UNIVERSITY OF SOUTHERN CALIFORNIA

SPACE TECHNOLOGY SUMMER INSTITUTE 1967

sponsored
by

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

COMPENDIUM OF NOTES

for the

1967 N.A.S.A. SUMMER INSTITUTE

Held at

The University of Southern California

June 19, 1967 to July 28, 1967

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

INTRODUCTION

It is clear that the space age is upon us. New spacecraft are launched almost daily. We have pictures of the moon's surface and soon plan to send men there. Quiet allusions to the explorations of our planets are heard and we have already sent unmanned probes to gain knowledge of Mars and Venus, our closest neighbors. The National Space and Aeronautics Administration has been intimately involved in this work and has recognized the need to train engineers in some of the specific skills used in the design of spacecraft. In order to provide such training the N.A.S.A. sponsored Summer Institute is being held in 1967 at the University of Southern California. Previously the Institute has been held at the California Institute of Technology, the University of California at Los Angeles, and at the University of Southern California.

The material presented here covers four major areas. The spacecraft Functional Design section presents a survey of the hardware design oriented subjects necessary for structural and thermal design of spacecraft as well as an introduction to propulsion. The Space Guidance and Control chapter provides the techniques required for guiding and controlling a vehicle in space. The Spacecraft Systems Engineering section presents the skills required to integrate an overall mission and to provide a complete vehicle capable of surviving its space environment and yet be capable of being built and flown. Finally the chapter on Space Communications presents the methods used to handle, store, and communicate data produced by the spacecraft.

PREFACE

The material presented here covers the major portion of class work for the 1967 N.A.S.A. Summer Institute held at the University of Southern California. Each chapter has been written by a different author. Some editorial changes have been made to allow consistent numbering of figures, tables, and equations; however, no attempt has been made to eliminate redundant notation.

The material is aimed at senior-level engineering students and will be presented in 120 classroom hours. In addition to covering the information presented here, the students will participate in a design group to practice some of the skills introduced in the classroom sessions.

ACKNOWLEDGMENT

The Institute is sponsored by the National Aeronautics and Space Administration under Contract No. NSR 05-018-055. The continuing activity of Dr. R. Edwards, who acted as Director to the Institute in 1966 is gratefully acknowledged.

I SPACECRAFT FUNCTIONAL DESIGN (R.S. HICKMAN)

In this section we will discuss some of the technical disciplines required to design a spacecraft, launch it, and keep it functioning. We will discuss rocket propulsion, vehicle structural analysis, and temperature control.

In addition we will discuss the structure of a planetary atmosphere since all spacecraft must at least begin their function in an atmosphere.

1 ROCKET PROPULSION

1.1 Basic Concepts

A rocket is simply a device which expels mass from a vehicle in a given direction with a relative speed so that thrust is achieved. It is therefore a system which is transient in that it devours itself to gain velocity.

Consider a vehicle in free space with a mass m(t) which moves in a straight line with speed v(t), and expels mass backward at a rate $\dot{m}(t)$ with a relative exit speed c. With no external forces acting on the vehicle, the momentum of the vehicle and the expended fuel will be independent of time and, if they started at rest, will be zero. Consider an expenditure of fuel, δm, from the vehicle to the exhaust. The net change of momentum must be zero.

Figure 1.1

From Figure 1.1 we write

$$\text{momentum at } t = \text{momentum at } t + \delta t$$

$$(v + \delta v)(m + \delta m) - \delta m(v - c) = mv \qquad (1.1)$$

Dividing by δt and allowing $\delta t \to 0$

$$m\dot{v} = -\dot{m}c = T \qquad (1.2)$$

where $\dot{v} = dv/dt$

The term on the right hand side of (1.2) is often referred to as the vacuum thrust.

In general, the speed c is a function of the propellant and nozzle design; in many applications it can be assumed to be independent of the burning

rate. With this assumption, Equation (1.2) can be integrated to yield:

$$v = v_o + c \ln m_o/m \tag{1.3}$$

Where the subscript o designates the initial conditions. High vehicle speeds can, in general, be achieved through high exit speeds and large ratios of initial to burnout mass.

We will find the concept of <u>impulse</u> to be useful. Impulse is defined as the time integral of a force and has the dimensions of lb sec.

Recall that in one dimensional Newtonian mechanics:

$$F = m\, dv/dt \tag{1.4}$$

where

F = force in lb

m = mass (# sec^2/ft)

so that

$$m(v - v_o) = \int_{t_o}^{t} F\, dt \tag{1.5}$$

And the impulse is a measure of the change in momentum.

In rocket engines, the <u>Specific Impulse</u>, I_{sp}, is defined as the impulse delivered divided by the weight lost during the thrusting period, i.e.,

$$I_{sp} = \int_{t_o}^{t} T\, dt/\text{WEIGHT OF PROPELLANT} \tag{1.6}$$

From Equation (1.2)

$$\int_{t_o}^{t} T\, dt = \int_{t_o}^{t} -\dot{m}c\, dt = c(m_o - m_f) = c_{mp} \tag{1.7}$$

consequently:

$$I_{sp} = \frac{c_{mp}}{wp} = \frac{c}{g} \tag{1.8}$$

where g is the acceleration of gravity at the surface of the earth. By definition I_{sp} has the dimension of seconds. It is clearly highest for fuels that leave the vehicle at the highest relative speed, c.

It should be emphasized that this is the "zero back pressure value" of specific impulse and is therefore only a function of the material. An effec-

tive I_{sp} could be defined by including backpressure and will in general be less than c/g.

The rocket designer is, among other things, charged with obtaining a design for which the corresponding I_{sp} is large. A good typical value for a chemical propellant rocket engine is 300 sec. We will examine the thermodynamic basisof this statement subsequently.

Calculation: Suppose a rocket engine has an I_{sp} of 300 sec, and a velocity change of 25,000 ft/sec is required, what is the minimum mass ratio attainable? From (1.3)

$$\ln \frac{m_o}{m} = \frac{25000}{9666}$$

$$\frac{m_o}{m} = 13.3$$

With this class of specific impulses, it is obvious that a very large initial mass will be required to inject a significant mass into orbit.

1.2 Thermodynamics and Gas Dynamics

We will consider the quasi-one dimensional steady form of the gas dynamic equations. The equations derived will be for conservation of mass, momentum, and energy. We shall consider only the inviscid form of the equations, and not consider dissipative effects. Much of the qualitative character of the flow may be obtained from these equations.

We consider the flow of an idealized gas with a streamtube area A which is a function of distance x. A streamtube is defined as that closed surface which is aligned with the velocity and forms a closed cylinder. The velocity u, and thermodynamic quantities, temperature T, and density, ρ, will be needed to characterize the flow.

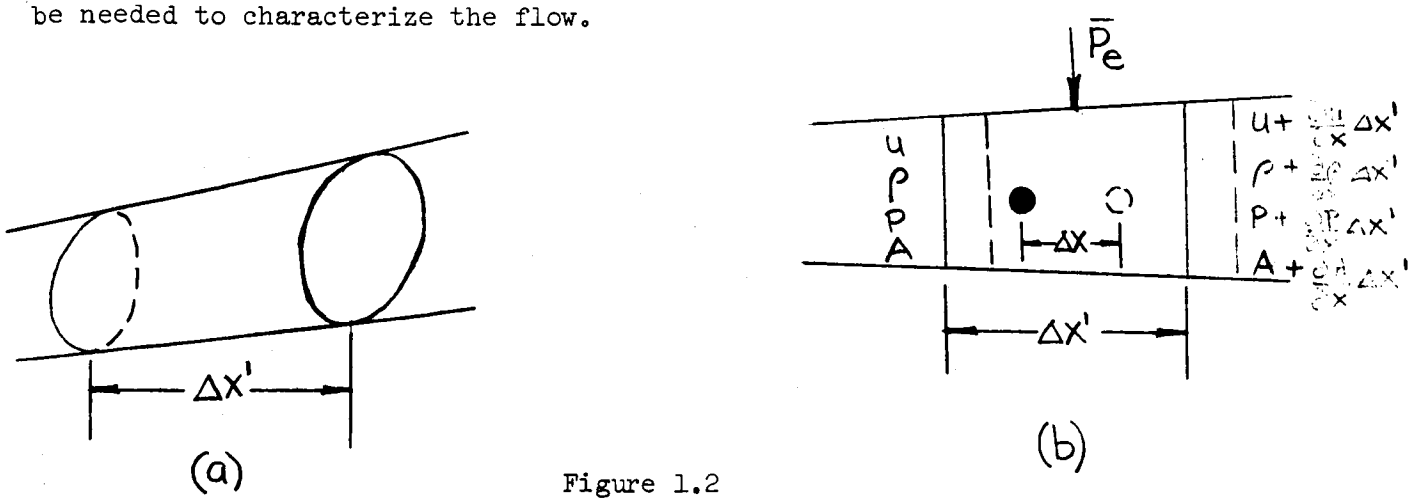


Figure 1.2

First we obtain the conservation of mass equation. From Figure 1.2 we can write

$$\dot{m}_{in} - \dot{m}_{out} = 0 \tag{1.9}$$

or

$$\rho u A = \text{CONSTANT} \tag{1.10}$$

We next obtain the conservation of momentum equation. Newton's law of motion says that

$$F = ma = m\, dv/dt \tag{1.11}$$

For the mass in Figure 1.2(b) m is approximately $\rho A \Delta x'$ neglecting terms like

$$\rho(\Delta x')^2 \frac{\partial A}{\partial x}$$

and

$$A(\Delta x')^2 \frac{\partial \rho}{\partial x}$$

The acceleration of the center of mass of the fluid is du/dt and is approximately

$$\frac{\Delta u}{\Delta t} = \frac{\Delta x}{\Delta t} \frac{\partial u}{\partial x}$$

The force F is just the sum of the pressure forces

$$- \frac{\partial P}{\partial x} A \Delta x'$$

Then using Equation (1.11)

$$- \frac{\partial P}{\partial x} A \Delta x' = (\rho A \Delta x') \frac{\partial u}{\partial x} \frac{\Delta x}{\Delta t} \tag{1.12}$$

Here it is important to observe that Δx is the positional shift of the center of mass of the fluid element during the time increment Δt. $\Delta x'$ is the length of the element and of course changes also with time; however, we are considering very short times so that this increase in length does not matter. Finally, we can clear this equation to get:

$$- \frac{\partial P}{\partial x} = \rho u \frac{\partial u}{\partial x} \tag{1.13}$$

where we recognize that u is

$$\lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t}$$

We finally obtain the conservation of energy equation. In order to discuss the energy equation we recall the form of the first law of thermodynamics which states:

$$\Delta Q = \Delta E + \Delta W + \Delta KE \tag{1.14}$$

Where ΔQ is the heat added to a system, ΔE is the internal energy change, ΔW is the work done, and ΔKE is the change in kinetic energy.

For the condition in Figure 1.2, we may again identify a fixed mass system; then during an increment in time Δt, the work is done at the faces of the elements which are normal to the streamline. The rate of work done on the element is PuA and the rate of work done by the element is

$$\left(P + \frac{\partial P}{\partial x}\Delta x'\right)\left(A + \frac{\partial A}{\partial x}\Delta x'\right)\left(u + \frac{\partial u}{\partial x}\Delta x'\right)$$

The net sum of these must equal the sum of the rates $\dot{Q}$, $\dot{KE}$, and $\dot{E}$.

Now again we must talk of the average of ΔKE and ΔE. We now limit ourselves to adiabatic flow for which $\dot{Q} = 0$. Then ΔE is

$$(\rho A \Delta x')\left(\frac{\partial e}{\partial x}\Delta x\right)$$

Similarly, the change in kinetic energy is

$$(\rho A \Delta x')\left(\frac{\partial ke}{\partial x}\Delta x\right)$$

Finally, we may write

$$PAu\Delta t - PAu\Delta t - \Delta t\,\Delta x'\,AP\frac{\partial u}{\partial x} - \Delta t\Delta x'\,Au\frac{\partial P}{\partial x} - \Delta t\Delta x'\,PA\frac{\partial A}{\partial x}$$

$$= \rho A \Delta x'\left\{\frac{\partial e}{\partial x}\Delta x + \frac{\partial ke}{\partial x}\Delta x\right\}$$

or

$$\Delta t\left(AP\frac{\partial u}{\partial x} + Au\frac{\partial P}{\partial x} + Pu\frac{\partial A}{\partial x}\right) - \rho A \Delta x\left\{\frac{\partial e}{\partial x} + \frac{\partial ke}{\partial x}\right\} = 0 \tag{1.15}$$

Now $\Delta x/\Delta t = u$ hence:

$$\frac{\partial e}{\partial x} + \frac{\partial ke}{\partial x} + \frac{P}{\rho}\left(\frac{1}{u}\frac{\partial u}{\partial x} + \frac{1}{A}\frac{\partial A}{\partial x}\right) + \frac{1}{\rho}\frac{\partial P}{\partial x} = 0 \tag{1.16}$$

Now using continuity (1.9) for the term in braces

$$\frac{\partial e}{\partial x} + \frac{\partial ke}{\partial x} - \frac{P}{\rho^2}\frac{\partial P}{\partial x} + \frac{1}{\rho}\frac{\partial \rho}{\partial x} = 0 \tag{1.17}$$

or

$$\frac{\partial}{\partial x}(e + k_e e_e + \frac{P}{\rho}) = 0 \tag{1.18}$$

The combination $e + P/\rho$ is defined as the enthalpy, h, and the energy equation becomes

$$\frac{d}{dx}(h + \frac{1}{2}u^2) = 0 \tag{1.19}$$

along streamtubes for adiabatic flow. Note that we have not restricted the state of the gas nor specified that it behaves perfectly (i.e., $P = \rho RT$ does not necessarily hold).

Equation (1.19) tells us how to calculate the absolute maximum value of I_{sp} since for a given nozzle system the maximum value of u will yield the maximum I_{sp}. For a streamline in a rocket motor (1.19) becomes

$$h + \frac{1}{2}u^2 = h_o \tag{1.20}$$

where h_o is the maximum stagnation enthalpy available in the motor. Clearly, the maximum attainable value of I_{sp} is reached when $h \to 0$. Then

$$u_{max} = g\, I_{spmax} = \sqrt{2h_o} \tag{1.21}$$

We now must relate h_o to the temperature in the rocket motor since we expect to be somehow limited in our structural design to temperatures in the neighborhood of 4000 to 5000°F.

To obtain a simple expression for h_o we will assume that the rocket exhaust gases are perfect and that the enthalpy is a function only of the temperature and is in fact just $C_p T_o$. This simple expression works well for monatomic gases and is justifiable for our use. For calculation for arbitrary gases we must evaluate h_o. Now the equation of state of a perfect gas is

$$P = \frac{\rho R_o T}{M} = \rho \bar{R}\, T \tag{1.22}$$

where R_o is a universal gas constant (1.98 BTU/# MOLE°R) and M is the molecular weight. R is the specific gas constant (R_o/M) and has units of BTU/#°R.

Now if C_p is the specific heat of the gas, the differential form of the first law says that

$$dQ\big|_{p=CONST} = de\big|_p + pdV\big|_p \tag{1.23}$$

or

$$C_p\, dT = C_V\, dT + R\, dT \tag{1.24}$$

or

$$C_p = C_V + R$$

Hence we see that for a perfect gas

$$C_p - C_V = R = \frac{R_o}{M} \tag{1.25}$$

Then returning to (1.21)

$$u_{max} = \sqrt{2h_o} = \sqrt{2C_p\, T_o} \tag{1.26}$$

Now the sound speed of a perfect gas is just $a = \sqrt{\gamma RT}$ where $\gamma = C_p/C_V$

hence

$$u_{max} = \sqrt{\frac{2C_p}{\gamma R}\,\gamma RT_o} \tag{1.27}$$

and

$$\frac{C_p}{\gamma R} = \frac{C_p}{\gamma(C_p - C_v)} = \frac{1}{\gamma - 1} \tag{1.28}$$

or

$$u_{max} = \sqrt{\frac{2}{\gamma-1}}\,\sqrt{\gamma RT_o} = \sqrt{\frac{2}{\gamma-1}}\sqrt{\gamma\,\frac{R_o}{M}\,T_o} \tag{1.29}$$

Now for a fixed gas and temperature γ is fixed. In fact, γ does not vary much ($\gamma = 1.67$ monatomic, $\gamma = 1.4$ diatomic, $\gamma = 1.2$ to 1.3 polyatomic) and for a fixed limit on T_o we see that u_{max} is largest for the lightest molecular weight. Of course in a real rocket the local flow enthalpy will not approach zero; however, it is possible to have h approach one or two per cent of $1/2\ u^2$ so that (1.26) is a good approximation for comparison of fuels.

1.3 Nozzle Flows

To actually obtain thrust from an engine we must use a shaped nozzle

to accelerate the high temperature gases to a high relative speed, u_{max}. We ask what must the shape of the nozzle be? To obtain the answer we start with (1.10). First differentiate the log of (1.10) with respect to x.

$$\frac{1}{\rho}\frac{d\rho}{dx} + \frac{1}{A}\frac{dA}{dx} + \frac{1}{u}\frac{du}{dx} \tag{1.30}$$

From thermodynamics we know that ρ is only a function of two other variables. Then

$$\frac{d\rho}{dx} = \frac{\partial\rho}{\partial s}\Big|_P \frac{ds}{dx} + \frac{\partial\rho}{\partial P}\Big|_s \frac{dP}{dx} \tag{1.31}$$

But

$$\frac{\partial P}{\partial\rho}\Big|_s = a^2$$

If a nozzle expands the flow adiabatically and reversibly then $ds \equiv 0$. Then

$$\frac{d\rho}{dx} = \frac{1}{a^2}\frac{dP}{dx} \tag{1.32}$$

and from (1.13)

$$\frac{1}{a^2}\frac{dP}{dx} = \frac{-\rho u}{a^2}\frac{du}{dx} \tag{1.33}$$

and (1.30) becomes

$$-\frac{u}{a^2}\frac{du}{dx} + \qquad + \frac{1}{u}\frac{du}{dx} + \frac{1}{A}\frac{dA}{dx} = 0 \tag{1.34}$$

Then if we define $M = u/a$ we see that

$$\frac{1}{u}\frac{du}{dx}\{1 - M^2\} = -\frac{1}{A}\frac{dA}{dx} \tag{1.35}$$

Since $u>0$ and $P/\rho a^2>0$ we see that for a subsonic flow for du/dx to be positive the nozzle must converge. At the point where $M = 1$ the nozzle area must pass through a minimum and then for a continued positive acceleration the area must increase when $M>1$.

Clearly, we wish to expand the hot rocket exhaust gases to as high a velocity as possible.

1.4 Energy Generation in Rockets

By far, the most common form for a rocket engine is the chemically driven rocket. These rockets are generally classed in two categories, solid and liquid. In the solid rocket, the fuel and oxidizer are cast together in a nearly solid rubberlike form. The liquid rocket system is usually a bi-propellant system in which the fuel is stored in one tank and the oxidizer in a second tank. The fuel and oxidizer are pumped, mixed, and injected into a combustion chamber where they are burned. There are advantages to each system, and each has its area of application.

The liquid engine has to date, achieved higher specific impulse than the solid engine and is easier to control in the sense that thrust levels may be varied by control of fuel flow rates, and shut off and restart are commonly achieved.

The solid engine has distinct advantages in simplicity (lack of plumbing) and general storeability, but it has disadvantages in that the thrust level cannot be easily controlled, the thrust level is often strongly influenced by temperature, cut off is at least more costly and restart is virtually impossible.

The physical characteristics of many propellants are listed in Glass tone, and will not be repeated here.

The primary objective is to achieve high specific impulse, or high combustion temperatures with low molecular weight. Hydrogen fuel and either oxygen or fluorine oxidizer yields specific impulses of the order of 400 sec.

Solid propellants as yet do not deliver specific impulse at this level. A figure of 300 sec would likely be very good. Many of the high I_{sp} rockets are loaded with metal which burns. The products of combustion then include particles of metallic oxide. The presence of these particles modifies nozzle design, so as to minimize thermal and momentum drag.

A solid propellant normally requires a given pressure level to sustain burning. Ignition is effected by firing an igniter which sets up the proper

pressure temperature environment. The burning rate (#/sec ft^2) can be correlated with pressure such that the rate is proportional to p^K, K>0. A high pressure environment then induces more rapid burning. In order to sustain a given thrust, the rocket grain should be designed so that the exposed area is roughly invariant with time. This, combined with heat transfer considerations is the reason for the star shaped cutout in some grain designs.

Various exotic schemes have been proposed. One, using a nuclear reactor, passes a nonreacting gas through a heat exchanger, which is heated by the nuclear reactor. A low molecular weight gas can be used and, at temperatures equivalent to those of chemical reactions, much higher specific impulses can be induced (velocity $\sim \sqrt{1/M}$). The weight of the reactor and shielding can be quite costly in performance and will probably make the nuclear rocket unuseable on vehicles at least as big as those used on current near earth missions.

Ion engines are devices which accelerate ions electrostatically. Very high exit speeds can be induced, but so far, only very low thrust levels have been achieved. There is a significant problem in space charge neutralization.

1.5 Optimum Operation of Rocket Systems

We will be concerned here with the question of optimum burning of fuel in order to achieve some objective. First, consider the problem of optimum thrusting of a vertically climbing rocket, with no aerodynamic drag, a constant specific impulse, and a constant gravitational field (flat earth).

$$m\dot{v} = -\dot{m}c - mg \tag{1.36}$$

or

$$\dot{v} = -\frac{\dot{m}}{m}c - g \tag{1.37}$$

here v

$$v = \frac{dy}{dt} \text{ and } \frac{dv}{dt} = v\frac{dv}{dy} \tag{1.38}$$

Then Equation (1.38) can be written in the form

$$\frac{d}{dy}\log m \exp(v/c) = -\frac{g}{cv} \tag{1.39}$$

The quantity log m exp(v/c) can be considered to be a measure of how fast a vehicle could go if its remaining fuel were expanded instantaneously (so that gravity impulse would have no effect). With such a program, log m exp(v/c) or log m + v/c would remain constant, and the maximum achievable speed would be given by

$$v_{max} = v + c \log \frac{m}{min} \tag{1.40}$$

The optimum thrust program may be shown to be that which maximizes log m exp(v/c) at every altitude since, the optimum could be exceeded by an impulsive burn otherwise. Obviously in this case, the optimum procedure is one of firing as rapidly as possible. The idea would be to fire instantaneously.

One can modify the program to account for aerodynamic drag by noting that, with drag:

$$m\dot{v} = - \dot{m}c - mg - \frac{1}{2} \rho v^2 S \, C_D \tag{1.41}$$

(here ρ is a function of altitude). Then, by similar manipulation,

$$\frac{d}{dy}(\log m \, e^{v/c}) = - \frac{g}{cv} - \frac{1}{2c} \frac{\rho v S C_D}{m} \tag{1.42}$$

Again, the optimum procedure minimizes the loss in m exp(v/c) at each station, and results in a velocity program such that,

$$\frac{\partial}{\partial v}\left(- \frac{g}{cv} - \frac{1}{2Cm} \rho C S C_D \right) = 0$$

For a constant C_D, one obtains:

$$\frac{g}{v^2} = \frac{1}{2m} \rho S C_D \tag{1.43}$$

This is a speed such that the drag equals the weight of the vehicle. Such a program will call for an impulsive burn until the drag equals the weight, followed by an acceleration such that the decrease in density is counteracted by the increase in speed such that the drag-weight balance is preserved.

2 STRUCTURAL DESIGN

2.1 General Behavior of Stress and Strain

The structural design of an object is one of the most important parts of an engineer's job. Automobiles, boats, and aircraft all must eventually be

translated into structural entities which bear the loads and do the job intended. Spacecraft are no different. We first define the system objective, then perform calculations to predict the system behavior and finally begin sizing parts and specifying materials. In this section we will approach the design of structures from a fairly basic point of view as far as stress and deflection calculations go. Our treatment of which material to finally select will not be as basic since many of the tradeoffs in material selection cannot be made entirely quantitative. We must, for instance, be concerned with combinations of ease of formation, magnetic properties, strength, rigidity, and dielectric properties.

We begin by defining stress. Stress is a force per unit area and has both magnitude and direction. Since in general the net stress at a surface may not be perpendicular to the surface, we define normal (tensile or compressive) and shear stress parallel to the surface. The stress is defined as

$$\vec{S} = \lim_{\Delta A \to 0} \frac{\vec{F}}{A} \tag{2.1}$$

where $\vec{F}$ is the resultant of the infinitisimal forces acting upon the area. The stress may, of course, be either purely torsional (shear) or axial (tension or compression). We consider a two-dimensional element shown in Figure 2.1.

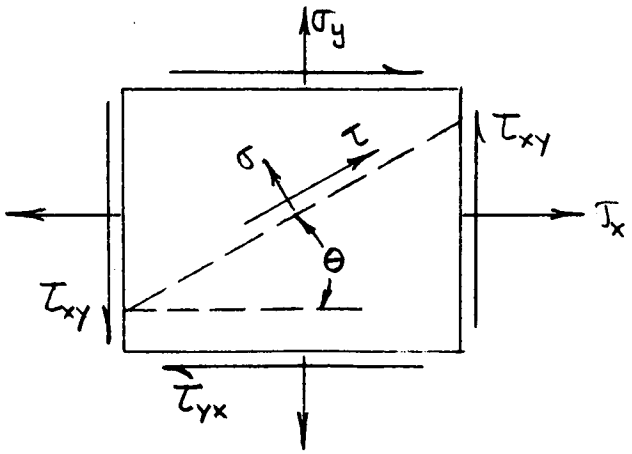


Figure 2.1

For no motion we find that $\tau_{xy} = \tau_{yx}$ and the stresses are as shown.

The general state of stress can be completely specified by Mohr's circle. To do this we must resolve the stresses according to the directions as shown. When this is done the stresses can be represented by a circle as shown in Figure 2.2.

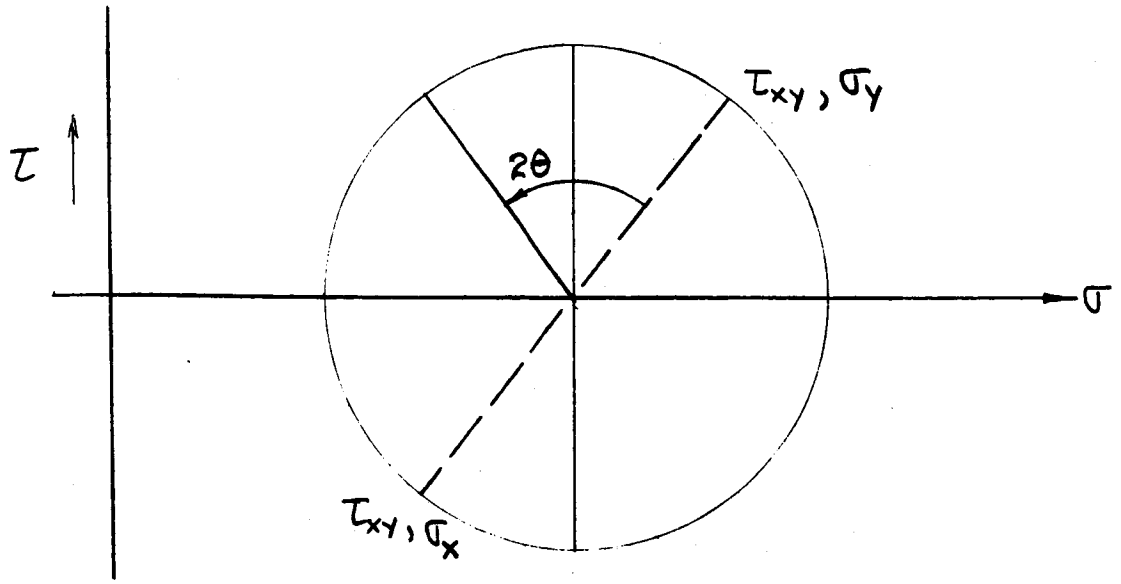


Figure 2.2

The significance of this is that both shear and normal stress exist in all loading conditions depending upon the direction of the plane of interest. Thus, if one has a material that fails at low levels of shear and the apparent loading produces normal stress only the part may still fail because of the existence of an angle at which the shear is high. As yet we have not attempted to relate the stress to strains in the material. The simplest relation, of course, is that the stress is linearly related to strain. Then

$$\sigma_x = E\varepsilon_x \tag{2.2}$$

where ε = strain = $\Delta\ell/\ell$ and E is of course the modulus of elasticity. The ability to induce stress in a direction normal to the direction of strain is defined by Poisson's ratio. Hence for (2.2)

$$\sigma_y = -m\varepsilon_x = \sigma_z \tag{2.3}$$

assuming an isotropic material. If m is small we can neglect (2.3) and treat only terms like (2.2).

2.2 Loads, Beams, Columns and Vessels

The simplest loading is pure tension or compression in a bar. Then

$$\sigma = \frac{F}{A} \tag{2.4}$$

Consider now a beam in which a moment must be carried. First we turn our attention to a two flange system. Then

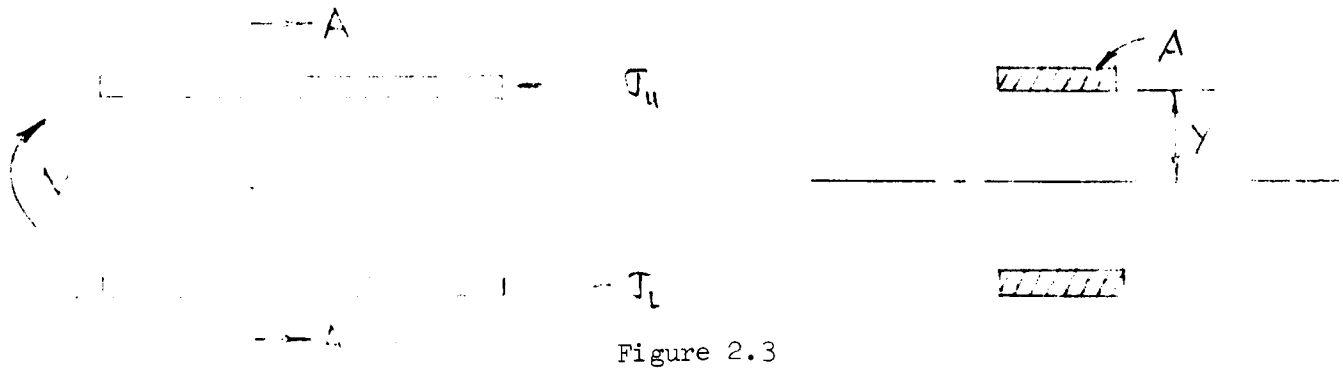


Figure 2.3

Clearly for

$$\Sigma F_x = 0$$

$$\sigma_u = -\sigma_L \tag{2.5}$$

and $2Ay\sigma = M$. Thus, a bending moment must produce stresses which balance one another and the moment. The continuous beam must meet the same requirements. For a beam carrying a moment the stress pattern in Figure 2.4 is found.

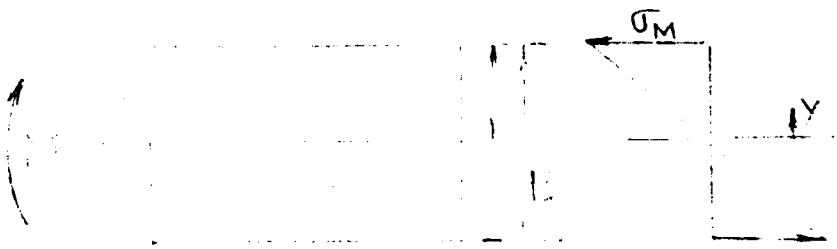


Figure 2.4

Here $\sigma = -\sigma_M y/C$ and

$$M = \frac{1}{C}\int_{-B/2}^{+B/2} \sigma_M y^2 W \, dy = \frac{\sigma_M}{C}\int_{-B/2}^{+B/2} y^2 W \, dy = \sigma_M I/C \tag{2.6}$$

The width, W, may be a function of y. The term I/C is a function only of the beam geometry and for most shapes is available in tables.

The static deflection of a beam is still unknown. We consider an elastic beam. If the beam is curved (due to stress) and if we consider the local strain we can write

$$\sigma = \varepsilon E \tag{2.7}$$

We still do not have enough information to specify the beam deflection. To finish our treatment we assume that initially plane sections remain plane after bending. Then

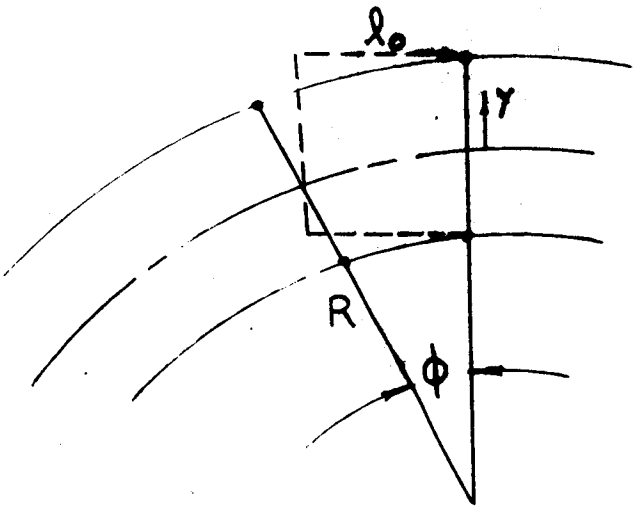


Figure 2.5

$$\varepsilon = \frac{\delta}{\ell_o} \tag{2.8}$$

where δ is the local elongation and

$$\frac{\delta}{y} = \frac{\ell_o}{R} \tag{2.9}$$

hence

$$\varepsilon = \frac{y}{R} \tag{2.10}$$

and

$$\sigma = E\frac{y}{R} = M\frac{y}{I} \tag{2.11}$$

Hence, we see that this result is comparable with the assumed stress distribution.

Now consider a length dx along the beam. The slope is dy/dx and is equal to $d\phi$ the angular displacement. If R is the radius of curvature of the beam we have $dx = Rd\phi$ and

$$\frac{d\phi}{dx} = \frac{1}{R} = \frac{M}{EI} \tag{2.12}$$

$$\Delta\phi = \Delta\left(\frac{dy}{dx}\right) = \int_{X_1}^{X_2} \frac{M}{EI}\, dx \tag{2.13}$$

For M = CONSTANT and $X_1 = 0$, $\phi_o = 0$

$$y = \frac{M}{EI}\frac{x^2}{2} \tag{2.14}$$

This is simply the equation for deflection of a simple beam clamped at one end and subjected to a constant couple, M. The above result does not help us in determining the shear distribution across a beam subjected to a shear load. In the sketch in Figure 2.6 we indicate a free body diagram for such a beam.

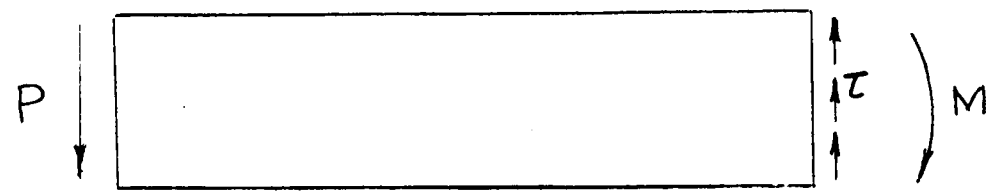


Figure 2.6

The shear load P must be resisted by the shear stress τ and a couple M must exist to maintain static equilibrium. We have already treated the stresses produced by the couple M. We now turn to the shear τ.

Since the upper and lower surfaces are free the shear stress must approach zero at these surfaces. Now even though we have accounted for the normal stresses produced by M we will find that M is important in the shear distribution. Consider an element dx long in a rectangular beam.

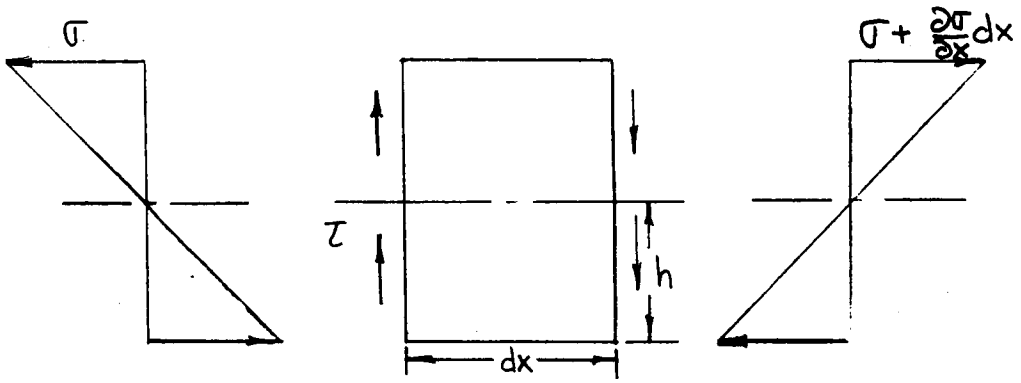


Figure 2.7

Now changes in the stress σ are due to changes in M due to increased distance from the applied load P. Since

$$\sigma = \frac{\sigma_m y}{h} = \frac{3}{2}\frac{M}{h^3}\frac{y}{w} \tag{2.15}$$

and

$$M = PX$$

$$\sigma = \frac{3}{2}\frac{P}{h^3 w} Xy \tag{2.16}$$

$$\frac{d\sigma}{dx} = \frac{3}{2}\frac{P}{h^3 w}\, y \tag{2.17}$$

Now consider a free body at y, Δy thick

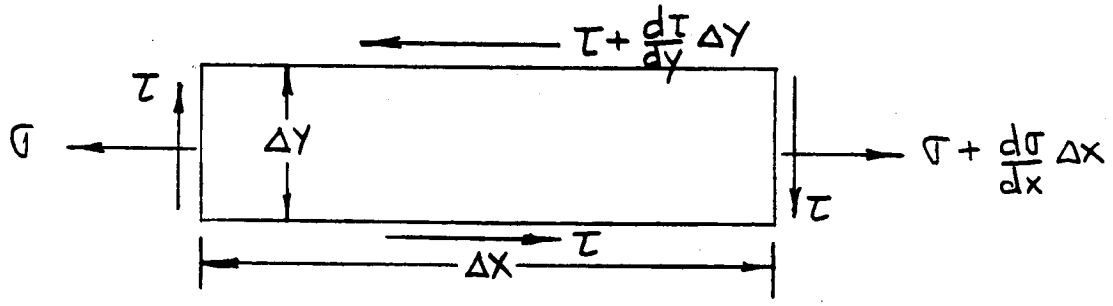


Figure 2.8

Now an X force balance yields (Note a y force balance shows that $\tau \neq \tau(X)$ at least to the approximation used here). Hence

$$\tau\Delta X - \tau\Delta X - \frac{d\tau}{dy}\Delta y\ \Delta X - \sigma\Delta y + \sigma\Delta y + \frac{d\sigma}{dx}\Delta X\ \Delta y = 0 \tag{2.18}$$

or

$$\frac{d\tau}{dy} = \frac{d\sigma}{dx} = \frac{3Py}{2h^3 w} \tag{2.19}$$

$$\tau = \frac{3P}{4h^3 w}\, y^2 + C \tag{2.20}$$

and at $y = \pm h$, $\tau = 0$ hence

$$\tau = \frac{-3P}{4wh}\left(\frac{y^2}{h^2} - 1\right) \tag{2.21}$$

and it is seen that τ is parabolic in distribution and

$$\int_{-h}^{+h} \tau\omega\ dy = P \tag{2.22}$$

We can obtain a formula for a general shape from (2.18) so that

$$\frac{d\tau}{dy} = \frac{d\sigma}{dx} \tag{2.23}$$

and σ can be related to M by the calculation

$$M = \int_{-h}^{+h} \sigma y \, W \, dy \tag{2.24}$$

plus the assumption that σ varies linearly with distance from the neutral axis.

We now turn to the structural behavior of colums. Spacecraft structures are commonly made up of trussed sections in which elements carry only axial loads. We are interested in the load which will collapse them. Clearly, if we exceed the structural limit in shear or compression, failure will occur.

There is another failure mode of importance. This is when the beam begins to deflect because of moments applied due to slight deflections caused by the axial load. From (2.12) we know that

$$\frac{d^2y}{dx^2} = \frac{M}{EI} \tag{2.25}$$

But if M is due only to the deflection y we have

$$\frac{d^2y}{dx^2} = - \frac{Py}{EI} \tag{2.26}$$

A solution to (2.26) is just

$$y = A \sin \alpha X \tag{2.27}$$

Now for a pinned column $y = 0$ at $X = 0$, $X = \ell$. Then

$$\alpha \ell = \pi \tag{2.28}$$

hence

$$\alpha = \frac{\pi}{\ell} \tag{2.29}$$

Substitution of (2.28) in (2.26) yields

$$- \frac{P}{EI} = - \frac{\pi^2}{\ell^2}$$

or

$$P = \pi^2 \frac{EI}{\ell^2} \tag{2.30}$$

The load defined by (2.30) must be applied for buckling to occur. This interesting form of failure can occur well below the elastic limit and is in fact an elastic failure. Of course, if the load P is sufficient to cause ordinary fail-

ure, buckling will not occur.

The buckling problem is one of the major problems faced by a rocket engineer. For maximum payload capacity he must make the structure as light and thin as possible; however, if he makes it too thin, then the acceleration load which must be transmitted from the heavy (hopefully) payload down to the rocket motor will buckle the vehicle. Some of the early attempts at launching the Vanguard rocket exhibited a severe buckling once the rocket got misaligned with the vertical and a slight additional moment due to gravity was applied to the vehicle.

The designers of the Atlas rocket solved this buckling problem ingeniously. They made the skin of the rocket very thin and left out all large structural members. Then they took this stainless steel balloon and pressurized it until it could sustain the vertical acceleration. In effect, they stiffened the container by placing a positive pressure in it. In later military and N.A.S.A. designs this technique was abandoned in favor of a rigid structure which would stand without pressurization.

So far we have discussed only idealized static systems. Idealized because we have ignored localized stresses caused, say, by application of a point load. The approximation that this localized effect can be ignored is referred to as St. Venant's principle. In addition to the idealization, the static treatment is open to criticism. This is becuase any loading process must begin and end. The transient behavior of the system must be characterized by some relaxation time, τ. Now is we are interested in times less than τ, we must include the dynamics. If our interest is primarily in times much longer than τ, we can treat the problem as if it were static.

If we somehow manage to apply and remove a load at frequency near $1/\tau$ we encounter resonance phenomena in which large amounts of energy may be stored in dynamic and elastic movement of the system. In such systems excessive loads can be generated which may cause failure. We could write simplified analyses which treat these dyanmics, but we will not since the analyses are fairly standard in dynamics texts and are not of excessive use. In practice, the following is done. First, for a given launch vehicle a recording is made of the motions of some surface during launch. Then this recording is played back through a suitably powerful speaker. The part to be tested is mounted to the speaker voice coil and is thereby subjected to a reasonable facsimile of the transient loading encountered by the spacecraft in the launch phase. Of course, such a

process cannot subject the vehicle to real sustained unidirectional acceleration loads; however, these are amenable to calculation and need not the same degree of scale testing.

3 THERMAL CONTROL

3.1 Introduction

We who have lived our lives in an atmosphere on the surface of the earth have been lulled into apathy toward solar radiation. Don't worry! If it is too hot this afternoon things will cool off a little tonight. When designing power stations the enormous amounts of energy which must be dumped to maintain the thermal cycle of the generation system are easily disposed of in a river or cooling tower.

Space does not provide these easy answers. For a constant mass vehicle in deep space the only way energy can enter or leave is via electromagnetic radiation. During an orbit transfer from earth to the moon a vehicle is subjected to a constant energy flux which must be reradiated away.

3.2 Radiant Energy

The simplest description of radiant energy is in the photon emission treatment. Here we assume a photon to be a massless particle which carries away energy. For a perfect emitter the law controlling emission is

$$q = \sigma T^4 \tag{3.1}$$

where

q = energy/unit area/unit time

and

T = Temperature (absolute!), $^\circ R$

with

$$\sigma = 0.1714 \times 10^{-8} \text{ BTU/hr ft}^2 \text{ }^\circ R^4$$

We postulate that for this radiant emission the intensity of radiation is constant in direction and depends only upon surface temperature. Then we specifically define I = intensity as radiant energy/unit solid angle/unit area along the ray of sight.

Then we observe that for a flat surface the net energy radiated per unit area from A_1 to A_2 is from Figure 3.1

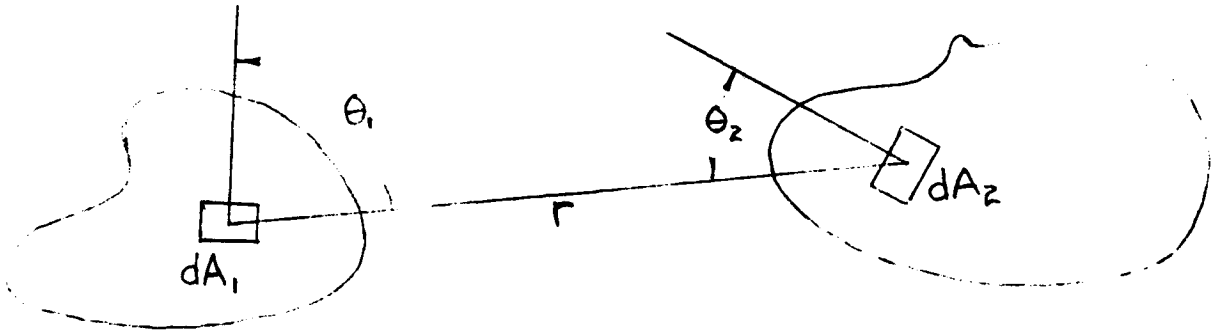


Figure 3.1

$$q_B = \int I \cos\theta_1 \, d\omega = \sigma T^4 \tag{3.2}$$

where

$$d\omega = \frac{dA_2 \cos\theta_2}{r^2} \tag{3.3}$$

Now

$$d\omega = \sin\theta \, d\phi \quad , \quad \phi = \text{azimuth angle}$$

hence for all radiation from a surface, assuming I is constant we have

$$I = \frac{\sigma T^4}{\pi} \tag{3.4}$$

Before proceeding we define a gray emitter. This is a surface which emits at a rate εq_B. The emissivity ε is in general a function of temperature. More specifically, we must define the wavelength dependent emissivity. All radiation can be characterized by a spectral distribution in wavelength. Then we define the spectral intensity per unit wavelength I_λ. Now

$$I = \int_0^\infty I_\lambda d_\lambda$$

Then consider ε to be a function of λ. We observe that if $\Delta\lambda I_\lambda$ falls upon a surface at equilibrium $I_\lambda \alpha_\lambda = \varepsilon_\lambda E_\lambda$ if $I_\lambda = E_\lambda$ we see that $\alpha_\lambda = \varepsilon_\lambda$. This is called Kirchoff's law. If the spectral distribution of energy falling upon a surface is the same as the spectral distribution of a black body at the surface temperature, then clearly $\bar{\alpha} = \bar{\varepsilon}$. In general, however, $\bar{\varepsilon} \neq \bar{\alpha}$ for a surface since the spectral source will be at a different temperature than the surface.

Now we wish to calculate the net transfer between two surfaces A_1 and A_2.

$$dQ_{12} = \bar{\alpha}_2 \frac{\sigma T_1^4}{\pi} \bar{\varepsilon}_1 \frac{dA_1 \cos \theta_1 \, dA_2 \cos \theta_2}{r^2} \tag{3.5}$$

Similarly

$$dQ_{2-1} = \bar{\alpha}_1 \frac{\sigma T_2^4}{\pi} \bar{\varepsilon}_2 \frac{dA_2 \cos \theta_2 \, dA_1 \cos \theta_1}{r^2} \tag{3.6}$$

The net flux is

$$dQ_{net} = (\bar{\varepsilon}_1 \bar{\alpha}_2 T_1^4 - \bar{\varepsilon}_2 \bar{\alpha}_1 T_2^4)\sigma \frac{\cos \theta_1 \cos \theta_2 \, dA_1 \, dA_2}{\pi r^2} \tag{3.7}$$

Now the local heat flux per unit area from area 1 to area 2 is

$$\int_{A_2} dQ = \dot{q} = (\bar{\varepsilon}_1 \bar{\alpha}_2 T_1^4 - \bar{\varepsilon}_2 \bar{\alpha}_1 T_2^4)\sigma \int_{A_2} \frac{\cos \theta_1 \cos \theta_2 \, dA_1 \, dA_2}{\pi r^2} \tag{3.8}$$

The integral is only a function of the geometry and is called the local configuration factor.

$$f = \int_{A_2} \frac{\cos \theta_1 \cos \theta_2 \, dA_2}{\pi r^2} \tag{3.9}$$

For simple geometry f can be calculated. The term

$$\frac{\cos \theta_2 \, dA_2}{r^2}$$

is just $d\omega$ the differential solid angle, hence

$$f = \int_{A_2} \frac{\cos \theta \, d\omega}{\pi} \tag{3.10}$$

Now we consider the reference hemisphere. A differential area on the sphere is $R^2 \, d\omega$. The projection of that differential

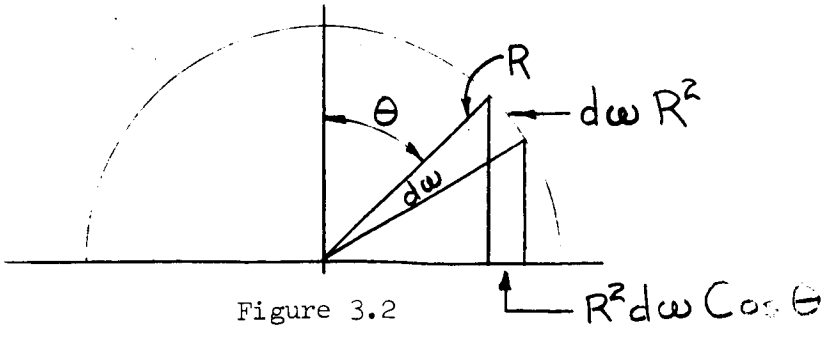


Figure 3.2

area on the base is $R^2\, d\omega \cos\theta$. Then if we can somehow project the area of A_2 onto a reference hemisphere and then project that area onto the base of the hemisphere the final projected area will be

$$A_P = \int_{A_2} R^2 \cos\theta \, d\omega$$

then f can be calculated as

$$f = \frac{Ap}{\pi R^2} \tag{3.11}$$

This formula is the basis for both experimental considerations and analytical calculations. For instance, consider a disc placed parallel to a surface. Take

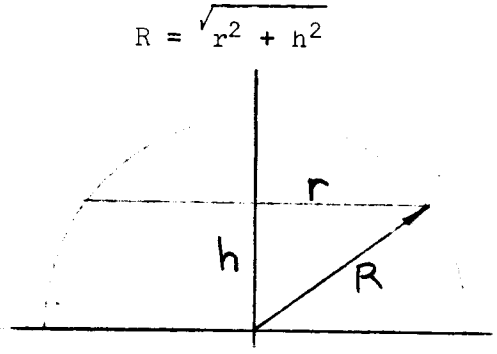


Figure 3.3

Then

$$\int_A R^2 \cos\theta \, d\omega = \pi r^2 \tag{3.12}$$

and

$$f = \frac{\pi r^2}{\pi(r^2 + h^2)} = \frac{r^2}{r^2 + h^2} \tag{3.13}$$

Now this case is easily calculated, but what if fifteen pipes of variable shape, distance, and angle are above the surface? Even if such a case could be calculated (in many cases it can be) the time required is prohibitive. Hence, we turn to the real usefulness of Equation (3.11). If a model is built and a point source of light is placed at the point from which f is desired then the shadow cast by any other surface onto a hemisphere centered at the point source provides a means of measurement of f. If a camera is placed far away, a photo will produce both the base of the hemisphere and the projected

area

$$\int_{A-2} \cos\theta \, d\omega$$

Figure 3.4 exhibits this technique.

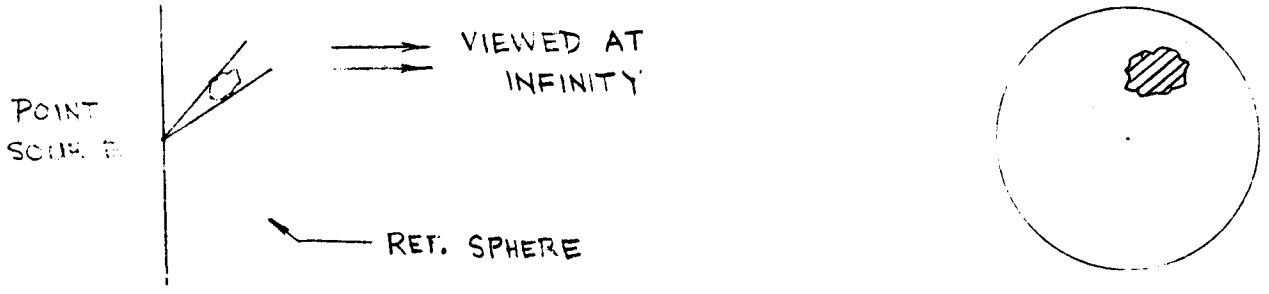


Figure 3.4

Unfortunately, it is extremely inconvenient to place a hemisphere over a model held at an arbitrary angle. To eliminate the sphere it is possible to do the same job with a parabolic mirror. Figure 3.5 demonstrates this technique.

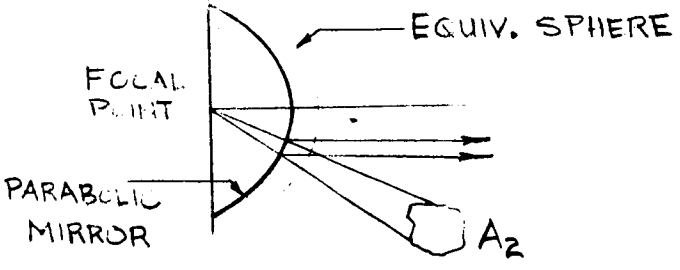


Figure 3.5

A parabolic mirror is made with its focal point located at the back center of the mirror. Any area outside the mirror may be viewed by aligning one's optical axis with the normal to the mirror. The areas seen can be related to the equivalent areas on a reference sphere as in Figure 3.4. Then a warped grid can be overlayed photos or a grid can be placed in the surface to directly evaluate f.

Once f is known as a function of position the overall configuration factor F is obtained from

$$F = \int_{A_1} f dA_1 \qquad (3.14)$$

This is achieved by obtaining f at different locations and finally calculating F.

Generally, we are only interested in F_{12} on a spacecraft to determine the configuration factor to black sky because this is our energy sink. Spacecraft tend to be at a constant temperature so that the net heat transfer between adjacent surfaces of a spacecraft by radiation is nearly zero. We can develop complex inter-area exchange calculations but they are seldom used because they simply are not important. Therefore, we are only really interested in F_{1B} or the configuration factor from the surface to black sky. If F_{1S} is the configuration factor from area 1 to the spacecraft, then

$$F_{1B} = 1 - F_{1S} \tag{3.15}$$

Hence, F_{1B} is the "left-over" configuration factor. The determination of F_{1B} remains a tedious job which must be performed for each important spacecraft surface.

The thermal balance is in general begun rather simply. Let q_g be the electrical energy generated and lost via resistance loads in a spacecraft. Further, let $\alpha_S\ A_S\ q_S$ be the energy received from the sun.
Then

$$\varepsilon\sigma\overline{A}\ \overline{T}^4 = q_g + \alpha_S\ A_S\ q_S \tag{3.16}$$

hence the overall average temperature can be calculated rather easily.

Each specific portion of the spacecraft, of course, can be analyzed separately and then conduction between regions accounted for. The spacecraft, like man (or because of his design philosophy), usually must be in a shirtsleeve (70°F) environment. If the spacecraft is a black ball ($\alpha = \varepsilon = 1$) then

$$\overline{T} = \left(\frac{q_g + \pi r^2\ q_S}{4\pi r^2\sigma}\right)^{1/4}$$

For

$$q_g = 0 \quad , \quad q_S = 454\ \text{BTU/ft}^3\ \text{hr}$$

$$T = \left(\frac{454 \times 10^8}{0.1714}\right)^{1/4} \approx 700°R = 240°F$$

This is hardly a shirtsleeve environment even when all electronic packages are off! The solution to this problem lies in several areas. First, we can make a material for which $\varepsilon \sim 1.0$ and $\alpha \sim 0.1$. This alone will drop T to 400°R, or -60°F. Second, we can make some portion of the spacecraft surface

quite hot, thus allowing the remainder to be cool. This technique is especially necessary when large amounts of power are to be generated on board the spacecraft. As a simple passive technique, one could place a disc some distance from a spacecraft. Then the disc could become red-hot and still not emit the same energy to the spacecraft as would be received from the sun.

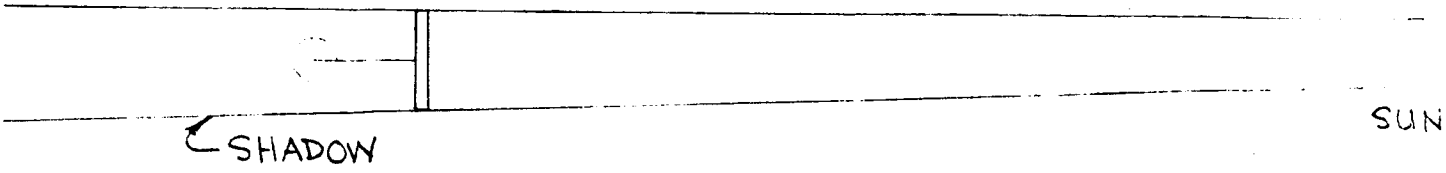


Figure 3.6

By a judicious choice of surface coatings and emitting surface location, we can control the temperature of components.

The design of unmanned vehicles is usually accomplished by treating the components as individual problems. The temperature control engineer is given the task of doing a complete thermal balance and specifying local coverings, active shades or thermal switches to maintain local temperature. We have not mentioned convection since it lacks primary importance in spacecraft thermal control. Conduction, however, is the dominant mode of thermal energy transfer within the spacecraft. This is because of the relative ability of solid material to conduct energy at low temperatures as compared to its ability to emit electromagnetic radiation. The major problem with conduction is the relatively unknown behavior of conductance joints in a high vacuum. All surfaces are microscopically rough and surface contact is at local points on the surface.

The best policy is to have clamped joints in as few locations as possible. Almost perfect conductance can be achieved by clamping gold or indium foil in critical conduction joints.

4 PLANETARY ATMOSPHERE

4.1 Density Distribution

The gaseous envelope surrounding a mass in space is called an atmosphere. It is primarily formed of material evolved by the planet. It remains on the planet because of gravitational attraction. Let us begin by considering the density distribution with distance from the planet's surface. The law of gravitation states that

$$F = G \frac{m_1 m_2}{r^2} \tag{4.1}$$

Where G is the universal gravitational constant, m_1, and m_2 are the masses of two objects, r is the distance separating them and F is the resultant force. For a finite amount of gas, m, distance r from the surface of the planet this force is just

$$F = mg \tag{4.2}$$

Where $g = g_o (r_o/r)^2$ and r_o is the radius of the planet. A force balance (Figure 4.1) then gives

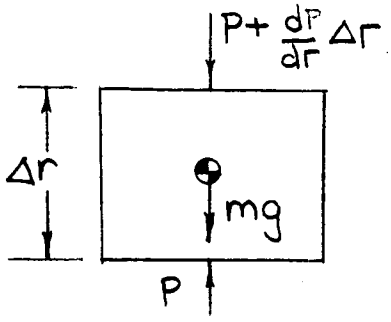


Figure 4.1

$$PA - PA - \frac{dP}{dr} \Delta r A = \rho \Delta r A g \tag{4.3}$$

then combining this with (4.2)

$$\frac{dP}{dr} = -\rho g = -\rho g_o \left(\frac{r_o}{r}\right)^2 \tag{4.4}$$

Now we must postulate an equation relating P and ρ. We observe that the earth's atmosphere is essentially at constant temperature. Hence, we use the ideal gas equation.

$$P = \rho RT \tag{4.5}$$

and assume that T is constant then

$$dP = RT d\rho \tag{4.6}$$

and

$$\frac{d\rho}{dr} = -\frac{\rho g_o}{RT} \left(\frac{r_o}{r}\right)^2 \tag{4.7}$$

Which integrates easily to yield

$$\frac{\rho}{\rho_o} = \exp -\left(\frac{g_o r_o}{RT}\left(\frac{r - r_o}{r}\right)\right) \tag{4.8}$$

and if we define $z = r - r_o$ and expect the fact that $z \ll r_o$ then

$$\frac{\rho}{\rho_o} = \exp -\left(\frac{g_o}{RT}\right) z \tag{4.9}$$

The term RT/g_o is sometimes called the scale height H. Then

$$\frac{\rho}{\rho_o} = \exp -\left(\frac{z}{H}\right) \tag{4.10}$$

The scale height depends upon the mass of the planet (through g_o) and upon the temperature. Since planets far away from the sun would be expected to be cooler than those closer to the sun, one would expect them to have a relatively higher rate of density decrease (because of a lower temperature).

Gases are composed of individual discrete molecules which move about most of the time as if there were no other molecules in the world. Since there are so many molecules involved in a unit mass of gas (about 10^{20}/in.^{3} at the earth's surface) we use statistical treatments to describe the velocity of the molecules. It turns out that even for a gas in which the apparent velocity is zero, individual molecules can have high velocity. At the fringes of a planetary atmosphere the gases do not collide often and if a gas has sufficient velocity it will escape the gravitational pull of the planet and will become lost. This is the apparent cause for the lack of atmosphere on the smaller planets. Even on the earth the light gas, helium, is lost.

4.2 Atmospheric Terminology

The various regions of the atmosphere are defined as follows:

Based on Temperature

Troposphere — The region nearest the surface, having a more or less uniform decrease of temperature with altitude. The nominal rate of temperature decrease is 6.5°K/km, but inversions are common. The troposphere, the domain of weather, is in convective equilibrium with the sun-warmed surface of the earth. The tropopause, which occurs at altitudes between 6 and 18 kilometers (higher and colder over the equator), is the domain of high winds and highest cirrus clouds.

Stratosphere — The region next above the troposphere and having a nominally constant temperature. The stratosphere is thicker over the poles, thinner or even nonexistent over the equator. Maximum of atmospheric ozone found near stratopause. Rare nacreous clouds also found near stratopause. Stratopause is at about 25 kilometers in middle latitudes. Stratospheric temperatures are of the order of arctic winter temperatures.

Mesosphere — The region of the first temperature maximum. The mesosphere lies above the stratosphere and below the major temperature minimum, which is found near 80 kilometers altitude and constitutes the mesopause. A relatively warm region between two cold regions; the region of disappearance of most meteors. The mesopause is found at altitudes of from 70 to 85 kilometers. Mesosphere is in radiative equilibrium between ultraviolet ozone heating by the upper fringe of ozone region and the infrared ozone and carbon dioxide cooling by radiation to space.

Thermosphere — The region of rising temperature above the major temperature minimum around 80 kilometers altitude. No upper altitude limit. The domain of the aurorae. Temperature rise at base of thermosphere attributed to too infrequent collisions among molecules to maintain thermodynamic equilibrium. The potentially enormous infrared radiative cooling by carbon dioxide is not actually realized owing to inadequate collisions.

Based on Composition

Homosphere — The region of substantially uniform composition, in the sense of constant mean molecular weight, from the surface upward. The homopause is found at altitudes between 80 and 100 kilometers. The composition changes here primarily because of dissociation of oxygen. Mean molecular weight decreases accordingly. The ozonosphere, having its peak concentration near stratopause

altitude does not change the mean molecular weight of the atmosphere significantly.

Heterosphere — The region of significantly varying composition above the homosphere and extending indefinitely outward. The "molecular weight" of air diminishes from 29 at about 90 kilometers to 16 at about 500 kilometers. Well above the level of oxygen dissociation, nitrogen begins to dissociate and diffusive separation (lighter atoms and molecules rising to the top) sets in.

Based on Ionization

Ionosphere — The region of sufficiently large electron density to affect radio communication. However, only about one molecule in 1000 in the F_2 region to one in 100,000,000 in the D region is ionized. The bottom of the ionosphere rises to 100 kilometers. The top of the ionosphere is not well defined but has often been taken as about 400 kilometers. The recent extension upward to 1000 km based on satellite and rocket data is shown.

The planetary atmosphere is formed by the evaporation of gases from the surface and is held by the gravitational force of the planetary mass. This explains the fact that the density of the atmosphere and its vertical distribution depend on the mass of a planet. Additional factors involved in the formation of a planetary atmosphere are: the chemical reaction between the component gases and the crust, the escape gases, and the heating and photochemical processes caused by solar radiation entering the atmosphere.

The sun is the principal source of energy causing atmospheric processes. All other sources of energy are negligible. Solar energy is in part reflected, in part absorbed or transmitted by the atmosphere. A portion of the radiative energy absorbed by the atmosphere is reemitted again and either escapes into space or reaches the surface of the Earth.

Solar radiation either directly or indirectly reaches the earth. Some energy consists of both direct and scattered solar radiation; some energy is reemitted by the atmosphere.

4.3 Absorption Reflection and Emission of Radiation

Radiant energy can be absorbed in a medium, reflected from a medium

or transmitted through a medium. Let us denote the ratios of the absorbed, reflected and transmitted energy to the incident radiation by a_λ , r_λ and τ_λ; the following relationship exists between these ratios.

$$a_\lambda + r_\lambda + \tau_\lambda = 1 \tag{4.11}$$

A mass element of the atmospheric gas dm emits in all directions the same amount of energy dE

$$dE_\lambda = e_\lambda \, dm d\omega d\lambda \tag{4.12}$$

where e_λ = mass emission coefficient. The total emission coefficient (as before) in all wave lengths is:

$$e = \int_0^\infty e_\lambda d_\lambda \tag{4.13}$$

The narrow energy beam of intensity I_λ passes through a medium of density ρ along the path dx. Absorption is proportional to the path length, density of the medium and intensity of the incident radiation

$$dI_\lambda = - k_\lambda \rho \, I_\lambda \, dx \tag{4.14}$$

where K_λ = the absorption coefficient given in $cm^2 g^{-1}$ and density ρ is given in g cm^{-3}. By integrating Equation (4.14) we get

$$I_\lambda = I_{\lambda o} \exp \left(-k_\lambda \int_o^x \rho dx \right) \tag{4.15}$$

This equation is known as Beer's Law. The quantity

$$u = \int_o^x \rho dx \tag{4.16}$$

is called the optical thickness.

As mentioned before, a portion of the incident radiation is reflected back, this amount, depending upon the reflective property of the body. Diffuse reflection is called albedo, and it is defined as the ratio of the reflected flux to the incident flux. It should be pointed out that reflection is also a function of wave lengths.

4.4 Law of Emission

Planck wrote the equation for the intensity of black body radiation in

the form,

$$I_{\lambda T} = \frac{C_1}{\lambda^5} \left[\exp \left(\frac{C_2}{\lambda T} \right) - 1 \right]^{-1} \qquad (4.17)$$

($c_1 = 2\pi c^2 h$, where c = velocity of light, h = 6.625×10^{-27} erg sec = Planck's constant; $c_2 = c\ h/k$, where k = 1.38×10^{-16} erg deg^{-1} = Boltzmann's constant). Experimentally it was found that $c_1 = 3.74 \times 10^{-5}$ erg/cm^2 xec and $c_2 = 1.430$c, deg, while λ_{1S} in cm and T om K°.

Inspection of Equation (4.18) demonstrates that the intensity increases very rapidly with rising temperature and that its peak emission shifts toward shorter wave lengths.

The total radiation intensity of a black body for all wavelengths will be found as,

$$F_{\lambda T} = \pi \int_0^\infty I_{\lambda T}\, d_\lambda = \frac{2\pi^5 K^4}{15c^2h^2} T^4 = \sigma T^4 \qquad (4.18)$$

This relation is known as the Stefan-Boltzmann equation. Using the Stefan-Boltzmann constant, we find $\sigma = 5.669 \times 10^{-12}$ watt cm^{-2} deg^{-4} = 0.814×10^{-14} cal cm^{-2} min^{-1} deg^{-4}; 0.1714×10^{-8} BTU/hr°R^4. In order to find the wave length of maximum intensity, we differentiate Planck's law with respect to the wave length and we equate to zero. From this relationship, Wien's displacement law is obtained.

$$\lambda_M = \alpha T^{-1} \qquad (4.19)$$

where $\alpha = 0.288$ cm deg. This equation also makes it possible to determine the color temperature of the body corresponding to maximum emission in certain wave lengths.

4.5 Solar Radiation in the Upper Atmosphere

The sun can be approximated by a black body radiating energy at a temperature of 6000°K. Figure 4.2 shows the spectral energy distribution of electromagnetic radiation computed for temperatures of 6000°K and 5700°K (solid line) and observed (dashed line). Short wave solar radiation reaches the earth; the earth in turn re-emits longer wave radiation (infrared), which is absorbed by the atmosphere, and in particular by H_2O and O_2; O_3, N_2O, and CH_4 absorb infrared radiation to a lesser extent. This absorption heats the atmosphere, which re-radiates the energy, a part of it downward; this energy provides addi-

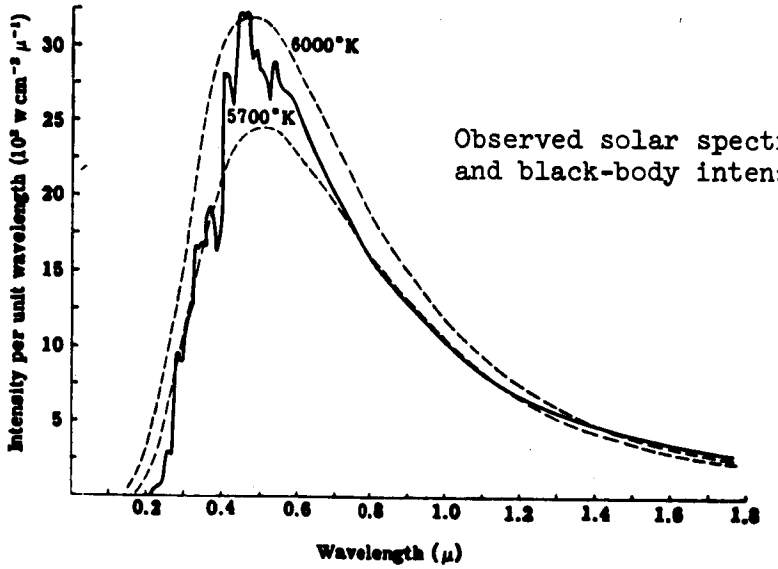


Figure 4.2

tional thermal energy for the earth's surface. This process of trapping solar energy is called the "greenhouse effect," a process in which infrared wave radiation prevails.

It should be noted that the radiation in the range between 8 and 12 microns can escape into space--with the exception of the ozone absorption hand at 9.6 microns. This wave interval is called the atmospheric infrared "window."

The density of the atmosphere diminishes with increase in altitude and accordingly there is a decrease in the total number of air molecules, water vapor, carbon dioxide and aerosol. This decrease in molecules changes the transmissivity of the atmosphere, and as a result there is a variation in the distribution of the solar spectrum at various altitudes.

The constituents of the upper atmosphere have no absorption spectra in the infrared; they are affected by short wave radiation, both near and extreme ultraviolet (EUV) and X-rays. The ultraviolet, EUV and X-rays comprise only about 1 percent of the total energy spectrum of the sun. Since these rays have quanta of large cross sections of absorption, they are absorbed by monoatomic and diatomic particles which prevail in the upper atmosphere. These regions of solar energy spectrum are responsible for the dissociation and ionization of the upper atmosphere.

4.6 Upper Atmosphere Temperature

Figure 4.3 shows the distribution of temperature. The temperature

profile, up to 85 km is based on balloon and rocket measurement. Above 85 km, the temperature is obtained from rocket and satellite observations. It can be seen that at low altitudes the constant temperature assumption is not too bad.

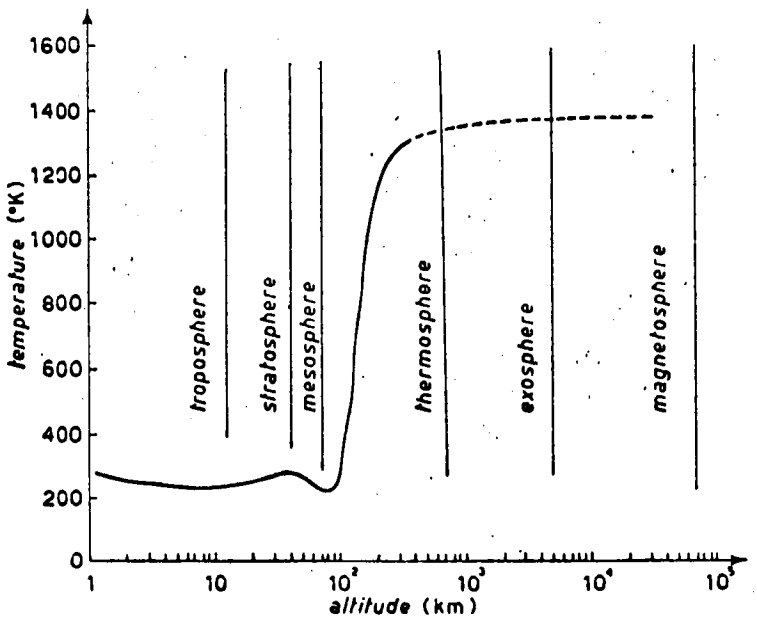


Figure 4.3. The Temperature Profile of the Earth's Atmosphere

As a result of the absorption by ozone (O_3) of solar ultra-violet radiation in the wave length range 2000A - 3000A, the temperature rises. This explains in Figure 4.3 the temperature increase from the sea level till 50 km height. At altitudes between 50 km and 90 km, the temperature decreases to 200° because of emission of infrared by carbon dioxide (CO_2) and diatomic oxygen (O_2). At higher altitudes the temperature rises again and at 300 km reaches the level of about 1200°K. This rapid temperature increase is caused by photo-dissociation and by photoionization of oxygen and nitrogen. These processes are produced by solar far-ultraviolet radiation.

II SPACE GUIDANCE AND CONTROL (N. NAHI)

1 LAPLACE TRANSFORMATION AND BLOCK DIAGRAMS

1.1 Introduction

The dynamical systems considered in this chapter can be represented by one or more ordinary linear differential equations with constant coefficients. While these equations can, in general, be solved by direct methods, application of the operational method of Laplace transformation has the following advantages:

a. The differentiation and integration operations are substituted by simple algebraic operations and consequently solution of the differential equation is reduced to a problem in algebra. The solution is in the operational form and tables are available to transform these solutions to the time domain. Many of the very important properties of the time domain solution can be directly obtained from its operational form.

b. Boundary or initial conditions are automatically included.

c. It provides a simple block diagram representation of the system.

1.2 Laplace Transformation

Let a function of time $f(t)$ be piecewise continuous and of exponential order.* The Laplace transform of $f(t)$ is denoted by $\mathcal{L}f(t)$ or $F(s)$ and is given by

$$\mathcal{L}f(t) = F(s) = \int_0^\infty f(t)\, \exp(-st)\, dt \tag{1.1}$$

The variable s is a complex quantity of the form $\sigma + i\omega$. As an example, the transform of the unit step function $f(t) = u(t)$ is

$$\mathcal{L}u(t) = \int_0^\infty \exp(-st)\, dt = \frac{1}{s} \tag{1.2}$$

The Laplace transform of $f(t) = \exp(-\alpha t)$

$$\mathcal{L}\exp(-\alpha t) = \int_0^\infty \exp(-\alpha t)\, \exp(-st)\, dt = \frac{1}{s + \alpha} \tag{1.3}$$

---

*The function $f(t)$ is of exponential order if there exists a constant such that $\exp(-\alpha t)|f(t)|$ is bounded for all t larger than some finite T.

The following relations between f(t) and F(s) can be established by the direct application of Equation (1.1).

a. Linearity

$$\mathcal{L}[af(t)] = aF(s) \tag{1.4}$$

$$\mathcal{L}[a_1 f_1(t) + a_2 f_2(t)] = a_1 F_1(s) + a_2 F_2(s) \tag{1.5}$$

b. Translation in s domain

$$\mathcal{L}[\exp(-at)f(t)] = F(s + a) \tag{1.6}$$

c. Differentiation

$$\mathcal{L}[\frac{d}{dt} f(t)] = sF(s) - f(0^+) \tag{1.7}$$

$$\mathcal{L}[\frac{d^n}{dt^n} f(t)] = s^n F(s) - s^{n-1} f(0^+) - \cdots - \frac{df^{n-1}}{dt^{n-1}}(0^+) \tag{1.8}$$

d. Integration

$$\mathcal{L}[\int_0^t f(t)\, dt] = \frac{F(s)}{s} + \frac{\int_0^{0^+} f(t)\, dt}{s} \tag{1.9}$$

e. Final Value

$$\lim_{s\to 0} sF(s) = \lim_{t\to\infty} f(t) \tag{1.10}$$

f. Initial value

$$\lim_{s\to 0} sF(s) = \lim_{t\to\infty} f(t) \tag{1.11}$$

1.3 Application of Laplace Transformation to the Solution of Differential Equations; Transfer Function

Let us consider the following network

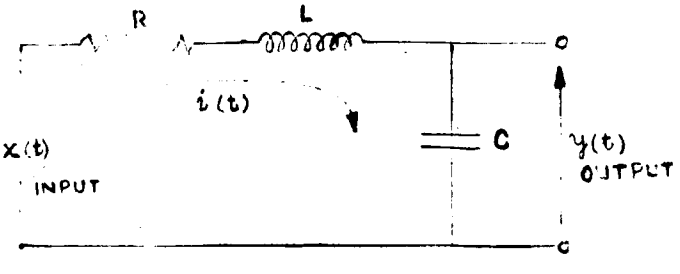


Figure 1.1

where $x(t)$ is the value of an input voltage source and $y(t)$ is the output of the network. The following equation can be written

$$x(t) = Ri(t) + L\frac{d}{dt}i(t) + \frac{1}{C}\int_0^t i(t)\,dt \tag{1.12}$$

where

$$\frac{1}{C}\int_0^t i(t)\,dt = y(t) \tag{1.13}$$

From Equation (1.13)

$$\frac{d}{dt}y(t) = \frac{1}{C}i(t) \tag{1.14}$$

$$\frac{d^2}{dt^2}y(t) = \frac{1}{C}\frac{d}{dt}i(t) \tag{1.15}$$

Substituting (1.13) - (1.15) into (1.12) yields

$$x(t) = LC\frac{d^2}{dt^2}y(t) + RC\frac{d}{dt}y(t) + y(t) \tag{1.16}$$

(1.16) is a differential equation relating $x(t)$ and $y(t)$. Assuming zero initial conditions, i.e., $y(0) = 0$; $d/dt\ y(t)\big|_{t=0} = 0$ and taking transform of both sides, we have

$$X(s) = LC\,s^2\,Y(s) + RC\,s\,Y(s) + Y(s) \tag{1.17}$$

The transfer function is defined to be the ratio of transform of the output to transform of input and is denoted by $G(s)$.

$$G(s) = \frac{Y(s)}{X(s)} \tag{1.18}$$

from (1.17)

$$G(s) = \frac{1}{LC\,s^2 + RC\,s + 1} \tag{1.19}$$

Hence

$$Y(s) = G(s)\,X(s) \tag{1.20}$$

Suppose it is desired to find the response to $x(t) = u(t)$, then

$$Y(s) = \frac{1}{LCs^2 + RCs + 1} \frac{1}{s} \tag{1.21}$$

which is the solution in operational form. If solution in time domain is required, the function y(t) which yields Equation (1.21) should be found. This process is referred to as inverse transformation. Standard tables are available which will facilitate the process of inverse transformation.

The representation of the input output relationships by means of Laplace transformation suggest a simple method of describing systems through block diagram notation. As an example, a single-degree-of-freedom stabilized platform is discussed below.

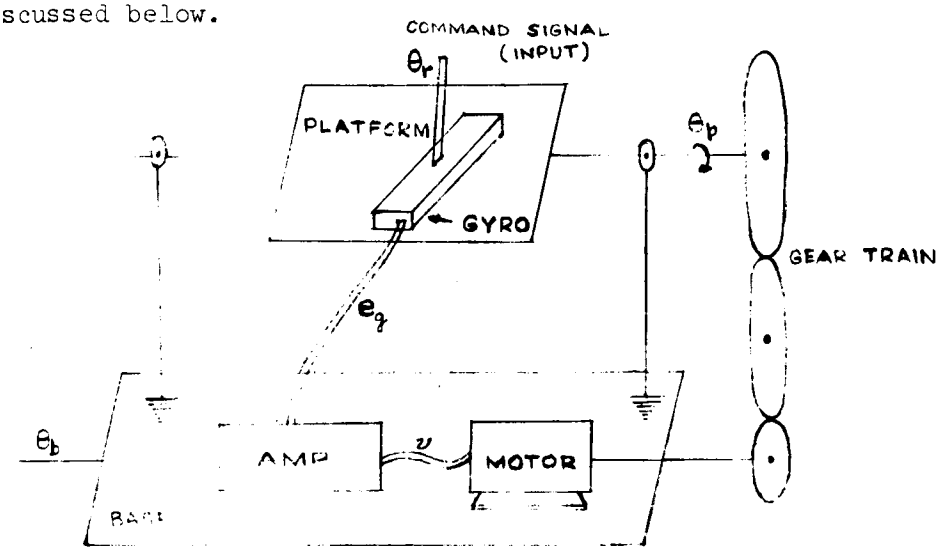


Figure 1.2. Single-Degree-of-Freedom-Stabilized Platform

The purpose of this control system is to keep the orientation of the platform fixed with respect to a reference for any angular movement of the base represented by θ_b. The following set of differential equations govern the dynamics of the system shown in Figure 1.2:

The gyro is an integrating gyro and produces a signal proportion to the difference between the desired platform angle θ_r and the actual angle θ_p

$$e_g = K_1(\theta_r - \theta_p) \tag{1.22}$$

where K_1 is the constant of proportionality. The amplifier amplifies the signal e_g by a factor K_2

$$v = K_2 e_g \tag{1.23}$$

The motor produces a torque f proportional to ν (it is assumed that f is unaffected by the motor shaft speed).

$$f = K_3 \nu \tag{1.24}$$

The torque f through the gear train with a gain K_4 causes the platform to rotate. The platform has an inertia J and damping B

$$K_4 f = J \ddot{\theta}_p + B \dot{\theta}_p \tag{1.25}$$

Taking Laplace transform of Equations (1.22) through (1.25) we have

$$E_g(s) = K_1 \left[\theta_r(s) - \theta_p(s)\right] \tag{1.26}$$

$$V(s) = K_2 E_g(s) \tag{1.27}$$

$$F(s) = K_3 V(s) \tag{1.28}$$

$$K_4 F(s) = J s^2 \theta_p(s) + B s \theta_p(s) = s \theta_p(s)(J S + B) \tag{1.29}$$

Now the following block diagram represents the set of Equations (1.26) through (1.29).

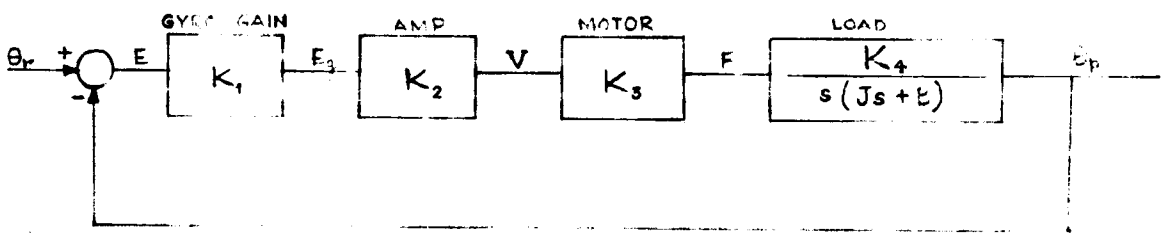


Figure 1.3. Block Diagram for Stabilized Platform

When $\theta_r = \theta_p$ then E = 0 and consequently F = 0 which in turn will cause no change in θ_p. The error E may be different from zero, either by a change in command θ_r or through a disturbance introduced by the motion of the base θ_b.

2 FEEDBACK SYSTEMS AND STABILITY

2.1 Open-Loop versus Closed-Loop

The function of stabilizing the platform in the example of the previous section can also be accomplished in the following manner. Referring to Figure 1.2, if a given change in θ_p is desired, then an appropriate signal ν can be applied to the motor to cause the right amount of change in θ_p. The system block diagram is in this case presented by Figure 2.1.

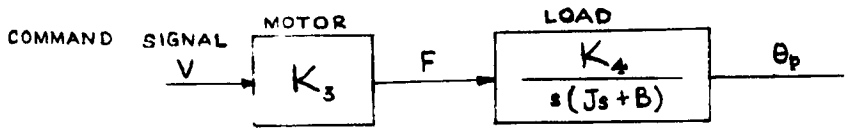


Figure 2.1. Open-Loop Control of θ_p

This type of control is called open-loop control, and requires precise knowledge of the constants K, J and B and is precise only if the base does not rotate after the initiation of command. In practice the parameters defining the dynamics of controlled objects either are not known with sufficient accuracy or their value may vary due to various conditions of the environment which are very difficult to account for. Consequently open-loop control is very sensitive to parameter variation and completely ineffective for eliminating any effect of disturbance inputs, such as motion of the base. Closed-loop systems, if designed properly, can reduce the effect of these problems to a point where the design is acceptable.

The following relationship between the transfer function of an open-loop system G(s) and that of a corresponding closed-loop system H(s) with a feedback transfer function K(s) is derived.

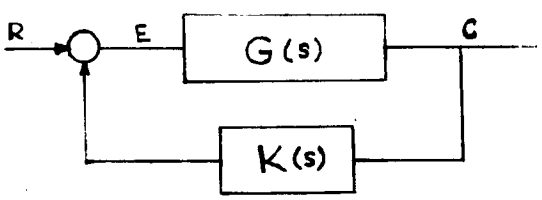


Figure 2.2

$$\frac{C(s)}{E(s)} = G(s) \tag{2.1}$$

$$\frac{C(s)}{R(s)} = H(s) \tag{2.2}$$

$$C(s) = E(s)\ G(s) = \left[R(s) - K(s)\ C(s)\right]\ G(S) \tag{2.3}$$

and finally, from (2.2) and 2.3)

$$H(s) = \frac{G(s)}{1 + K(s)\ G(s)} \tag{2.4}$$

2.2 Stability

A system is called stable if it returns to the "rest" condition after an initial perturbation. It is evident that this property is required of all practical control systems. In the stabilized platform example of Section 1.3, it means that, for example, in the absence of any change in command signal and base movement, if the platform is perturbed from its rest position, it returns to that condition automatically. Consequently, it is important to determine the stability condition of a system from the governing equations.

A differential equation describing a stable linear system has the property that its characteristic equation has roots with negative real parts. This guarantees that the homogeneous solution will have an exponentially decaying element ($e^{-\alpha t}$, $\alpha>0$) in each additive part, hence, all transients will decay out with time. The characteristic equation of a system is readily available as the denominator polynomial of the closed-loop transfer function. The stabilized platform example of Section 1.3 will be used here for illustration of stability analysis.

From block diagram of Figure 1.3, the closed-loop transfer function is given by

$$\frac{\theta_p(s)}{\theta_r(s)} = \frac{\dfrac{K}{s(J\,s + B)}}{1 + \dfrac{K}{s(J\,s + B)}}$$

$$K = K_1\, K_2\, K_3\, K_4 \tag{2.5}$$

Hence

$$\frac{\theta_p(s)}{\theta_r(s)} = \frac{K}{J\,s^2 + B\,s + K} \tag{2.6}$$

The differential equation relating input $\theta_r(t)$ to the output $\theta_p(t)$ can be deduced from (2.6) directly by identifying s as d/dt (the differential operator)

$$J\,\ddot{\theta}_p(t) + B\,\dot{\theta}_p(t) + K\,\theta_p(t) = K\,\theta_r(t) \tag{2.7}$$

which has the characteristic equation given by the denominator of (2.6)

$$J\,s^2 + B\,s + K = 0 \tag{2.8}$$

The polynomial (2.8) has the following two roots

$$-\frac{B}{2} \pm \sqrt{\frac{B^2}{4} - K J} \tag{2.9}$$

Since B, K and J are positive quantities, these two roots both have negative real parts for all values of these parameters. Notice that if feedback was positive instead of negative (i.e., for example, when the gear train is connected wrong) then (2.5) should be replaced by

$$\frac{\frac{K}{s(J s + B)}}{1 - \frac{K}{s(J s + B)}} \tag{2.10}$$

which yields the following two roots for the characteristic equations

$$-\frac{B}{2} \pm \sqrt{\frac{B^2}{4} + K J} \tag{2.11}$$

which clearly indicates that one of the roots is always a positive number.

In the area of classical control systems, a number of graphical and numerical techniques are available by which the stability of a linear system can be determined.

3 SPACE VEHICLE SYSTEMS

3.1 General Description of Mathematical Models for Various Space Vehicle Operations

The control functions of a space vehicle may be divided into three parts: a) attitude control, b) guidance, and c) tracking. In the following the above operations are described in brief.

a. Attitude Control - There are various reasons for the requirement that the orientation of a vehicle remain at a desired relationship, with respect to some reference set of axis. To name a few, a known orientation is needed when the vehicle is thrusted, a fixed orientation is needed when the vehicle is thrusted, a fixed orientation is desired when there are astronauts on board, a controlled orientation is necessary when an on-board camera (or any similar instrument) is to point in a specific direction. The reference system may be a fixed set of directions in space, which is obtained by means of an inertial platform or by direction of location of some stars. In case of an earth satellite, a reference system may have one axis which at any time passes through the center of the earth.

b. Guidance of Space Vehicles - Any space vehicle probably has a mission to perform, and that requires the transfer of the vehicle from one point in space to another. The requirement and conditions of any mission in general, adds more restrictions on how the transfer is to be accomplished and what trajectory to be followed. For example, it may be desired to land a vehicle on moon. A specific area on the surface of the moon may be chosen as the desired landing position. In order not to crash land, restrictions should be put on the terminal velocities of the vehicle. Furthermore, the value of the thrusts of the on-board engines are limited by a known quantity.

c. Tracking - Tracking is an important function to be carried out on board many future space vehicles. In case of rendezvous and docking of two vehicles, the required accuracies usually exceed those obtained by means of ground tracking of both vehicles. Future vehicles will definitely be required to be more autonomous (self governing) than those of the past. This appears in form of lessening necessity for reliance on information supplied by a ground base. Consequently, the necessary information for various pursuit and rendezvous missions have to be obtained on board by means of radar, optical or passive detection.

The above three functions will be treated in more detail in the next three sections. In the remainder of this section, some general properties of the mathematical models for the vehicle in the various modes of operation will be discussed.

In any of the desired operations, there is a part or all of the vehicle with dynamics which have to be controlled and there is mechanism for inducing the necessary control. In Figure 3.1 these are referred to as controlled object and the controller, respectively. The information available to the controller is the reference (or command) input or inputs and the output variable or variables which are fed back through the use of various sensors. For example in the case of tracking the controlled object is the antenna system, the controller is the drive (motor) which can re-orient the antenna and the sensor, may be a rate integrating gyroscope indicating the deviation of the center beam of the antenna from a reference axis. In the case of attitude control, the controlled object is identified by a set of equations representing the dynamics of the body of the

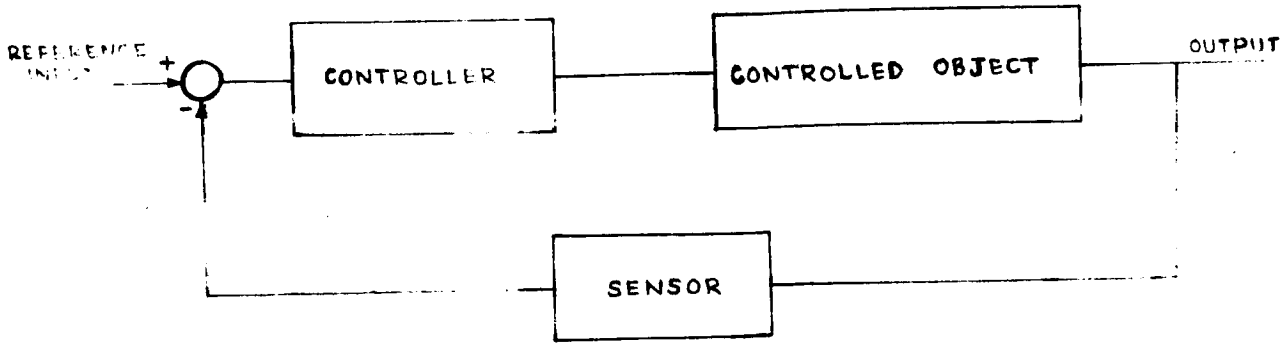


Figure 3.1. General Block Diagram of a Control System

vehicle, including the antenna system and any other object occupying the vehicle, with respect to a coordinate system. TXe controller may be a set of small reaction jets which when fired, will exert a torque causing a body reorientation. The sensors are a number of position and rate gyroscopes yielding the angular and angular rate deviations of the body with respect to a set of reference axis. Finally, in the case of guidance, the whole system including the antenna system and all the other instruments and astronauts can be assumed as a point mass which will represent the controlled object. The controller is one or more propulsion mechanisms which exert a vector valued force on the point mass (the controlled object). The sensors may be the tracking system on board which in turn will supply information concerning any deviation of the vehicle position, velocity and acceleration with respect to some desired values.

From the above discussion, it is seen that, for the same space vehicle, the controlled object may involve a part or whole of the vehicle depending on various functions to be performed.

3.2 Attitude Control of Space Vehicle Systems

The attitude control of a vehicle, i.e., maintaining a desired orientation for the body of the vehicle, is basically accomplished by two different procedures, passive and active. Passive attitude control is usually used for unmanned vehicles where a relatively sophisticated control of the orientation may not be necessary. The object here is to design a system which is stable in the vicinity of some desired orientation in other words, any perturbation about the desired orientation reduce to zero in time. The main sources of perturbing torques are: gravity gradient; atmospheric pressure; electromagnetic induction

and solar radiation. The methods of passive stabilization include: spin stabilization; balancing one perturbing torque against another; energy dissipation and tuned pendulum. The active attitude control is accomplished through the use of applied torques (e.g., generated by jets mounted on the body of the vehicle) in order to perform desired corrections in the orientation of the vehicle. Active attitude control is necessary when the mission calls for high accuracy and speed in response to a command input.

Let Figure 3.2 represent a space vehicle. The point o is the center of mass, ox, oy, oz are a set of orthogonal axis fixed on the body.

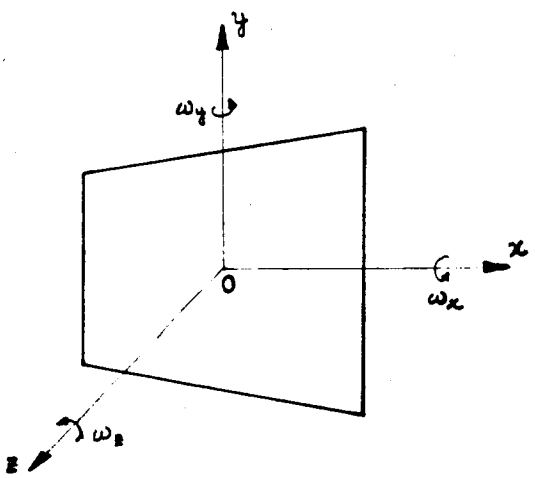


Figure 3.2. Space Vehicle

The angular velocities with respect to x, y, z axis are denoted by w_x, w_y, w_z. The moment of inertias are referred to as I_x, I_y, I_z. There may be one swiveling engine which induces a torque T with components T_x, T_y, T_z or a set of fixed, body mounted thrusters may produce these component values. The Newton's law of motion gives the following relationship.

Applied torque T = Rate of change of angular momentum with respect to an inertial coordinate system (3.1)

With respect to the body fixed axis x, y, z the law of motion assumes the well known Euler equations.

$$T_x = I_x \dot{\omega}_x + (I_z - I_y)\omega_g \omega_z \tag{3.2}$$

$$T_y = I_y \dot{\omega}_y + (I_x - I_z)\omega_z \omega_x \tag{3.3}$$

$$T_z = I_z \dot{\omega}_z + (I_y - I_z)\omega_x \omega_y \tag{3.4}$$

For a body with three axis of symmetry $I_x = I_y = I_z = I$ and consequently (3.2) to 3.4) become

$$T_x = I \dot{\omega}_x \tag{3.5}$$

$$T_y = I \dot{\omega}_y \tag{3.6}$$

$$T_z = I \dot{\omega}_z \tag{3.7}$$

In general, even when the above symmetry does not exist, the Equations (3.2) to (3.4) can be approximated by (3.5) to (3.7) since w_x, w_y, w_z are kept relatively small.

Since Equations (3.5) to (3.7) are completely independent of each other, only one such as the rotation about the x axis, is considered in the following. Let θ_x be the rotation angle about x axis with reprect to a reference and θ_{xd} be the desired value of the rotation about this axis. We have the following equations

$$T_x = I \dot{\omega}_x \tag{3.8}$$

$$\omega_x = \dot{\theta}_x \tag{3.9}$$

Furthermore, the thruster produces a torque proportion to the difference $\theta_{xd} - \theta_x$ with a constant of proportionality K. Therefore the following block diagram results.

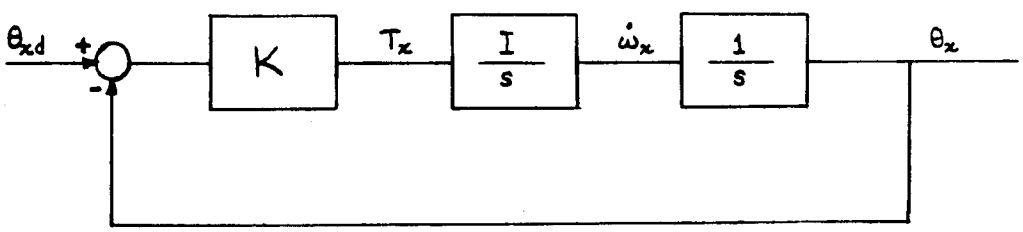


Figure 3.3. Block Diagram

The difference quantity $\theta_{xd} - \theta_x$ is the output of a rate integrating gyro with a reference axis θ_{xd}. The output Laplace transform is

$$\theta_x(s) = \theta_{xd}(s) \frac{\frac{KI}{s^2}}{1 + \frac{KI}{s^2}} = \theta_{xd}(s) \frac{KI}{s^2 + KI} \tag{3.10}$$

This represents an unstable system since the roots of the denominator polynomial do not have negative real parts. To see this better, let us assume we would like to turn the vehicle around the x axis by one unit. Then

$$\theta_x(s) = \frac{1}{s} \frac{KI}{s^2 + KI} = \frac{1}{s} - \frac{s}{s^2 + KI} \tag{3.11}$$

Consequently

$$\theta_x(t) = 1 - \cos \sqrt{KI}\, t \tag{3.12}$$

which clearly shows that $\theta_x(t)$ rather than approaching 1 will oscillate about the desired value. What is now needed to achieve the desired goal is referred to as a compensator. A very common way of stabilizing the above system is to add a rate feedback path. A rate gyroscope will produce an output proportion to angular rate $\dot{\omega}_x$. Let the constant of proportionality be K_1. The block diagram of Figure 3.3 is modified as

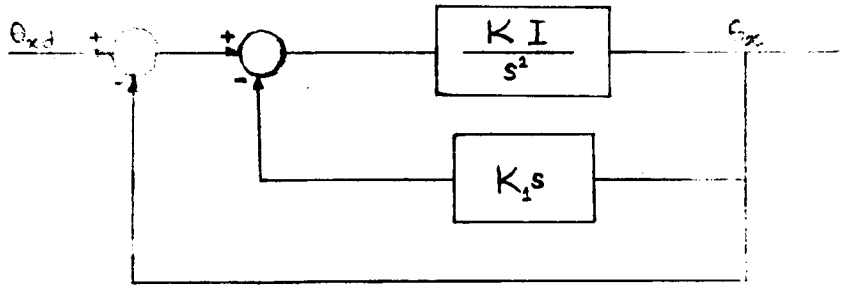


Figure 3.4. The Compensated System

The transform of the output is

$$\theta_x(s) = \theta_{xd}(s) \frac{\frac{KI}{s^2}}{1 + \frac{KI(1 + K_1 s)}{s^2}} = \theta_{xd}(s) \frac{KI}{s^2 + K K_1 I s + KI} \tag{3.13}$$

The system is now stable (the roots having negative real parts) and the solution to $\theta_{xd}(t) = 1$ is (for K_1 small)

$$\theta_x(t) = 1 + \alpha \exp\left(-\frac{K K_1 I}{2} t\right) \sin(\omega_o t + \phi)$$

$$\alpha = \frac{1}{\sqrt{1 - 1/4\ K_1^2\ KI}}$$

$$\omega_o = \sqrt{KI}\ \sqrt{1 - 1/4\ K_1^2\ KI}$$

$$\phi = \tan^{-1} \frac{\sqrt{1 - 1/4\ K_1^2\ KI}}{1/2\ K_1\ \sqrt{KI}} \qquad (3.14)$$

which clearly shows that as time is increased, the desired orientation will be approached.

In the above analysis, the operation of the thrusters were assumed linear, i.e., the output torque T_x was assumed to be proportional to its input, $\theta_{xd} - \theta_x$. Since reaction jets are commonly used in space vehicle applications (due to reliability and weight considerations) the operation is far from linear and is approximately presented by Figure 3.5 where e is the input to the torquer.

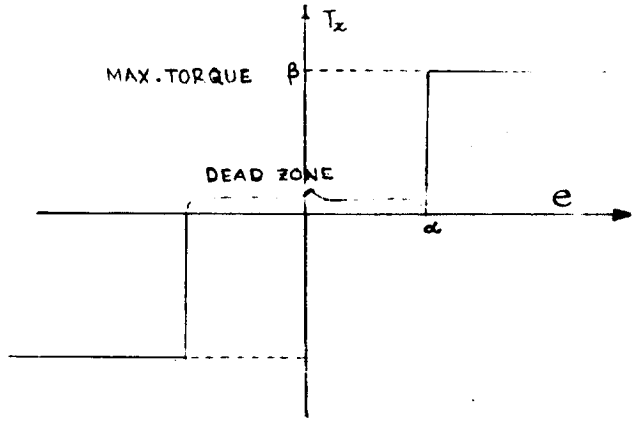


Figure 3.5. The Characteristic of Torquers

The block diagram of Figure 3.4 should be modified.

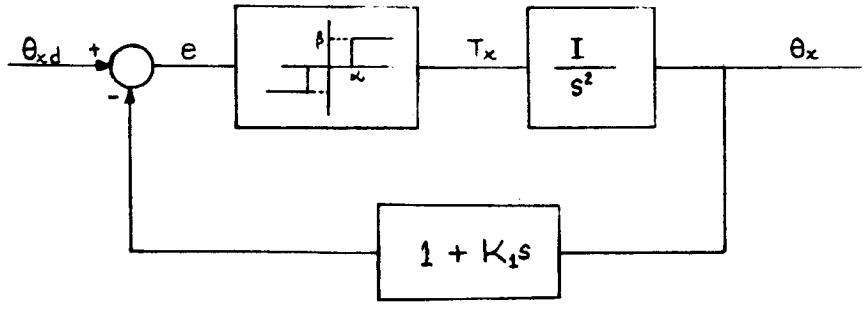


Figure 3.6. Nonlinear Attitude Control Loop

If $K_1 = 0$ (i.e., no rate feedback) it is easy to show that the output will oscillate about the desired input θ_{xd} (e.g., for θ_{xd} = constant). As an example, let the desired rotation θ_{xd} be unity and this is applied to the system as a step command. If $e<\alpha$ (2α is the width of the deadzone or inactive zone of the torquers) the system will not become active. If $e>\alpha$ then a torque $T_x = \beta$ will be applied to the vehicle which will yield (e(t) is the input to the nonlinearity)

$$\theta_x(t) = \frac{1}{2} I\beta t^2 \tag{3.15}$$

$$e(t) = 1 - \frac{1}{2} I\beta t^2 - K_1 I\beta t \tag{3.16}$$

This continues until e(t) is reduced to $+\alpha$ when the body will keep moving then on its own inertia until $e(t) = -\alpha$. At this time, a reverse torque will be applied. Time plot of this operation with and without K_1 term (i.e., K_1 = nonzero and $K_1 = 0$) will reveal that the system has a damped response for $K_1>0$ and oscillatory response for $K_1 = 0$. In order to verify this assertion, let us define an equivalent gain for the nonlinearity. Assume the input to the nonlinearity is a sine wave E sin wt. The output waveform is sketched in Figure 3.7.

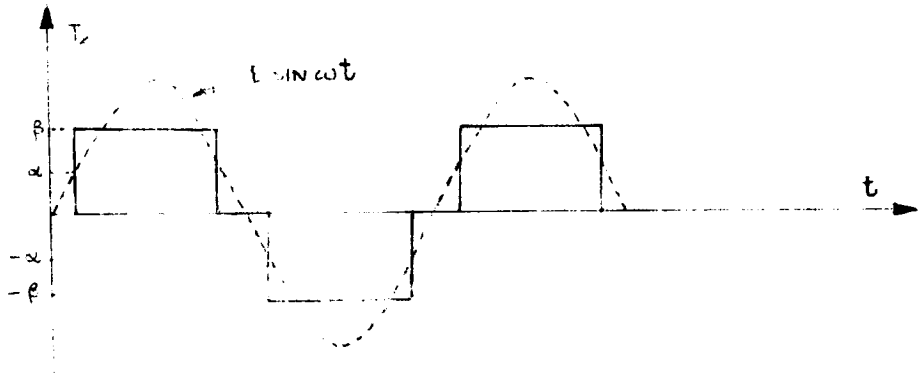


Figure 3.7. Response of Torquer to Sine Input

The fundamental component of the output waveform is obtained by a fourier series expansion.

Magnitude of Fundamental Comp. of

$$T_x = \frac{4\beta}{\pi}\sqrt{1 - \frac{\alpha^2}{E^2}} \qquad ; E > \alpha$$

$$= 0 \qquad ; E < \alpha \tag{3.17}$$

The ratio

$$\frac{4\beta}{\pi}\sqrt{\frac{1-\frac{\alpha^2}{E^2}}{E}}$$

is referred to as the describing function of the nonlinearity.

The block diagram of Figure 3.6 can be represented as

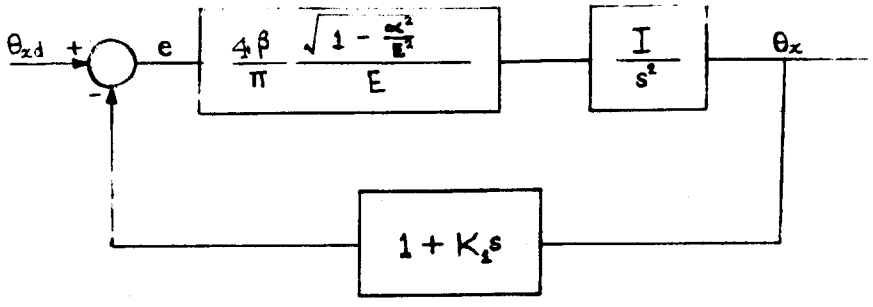


Figure 3.8. Block Diagram with Describing Function

Now let us conjecture that the loop is not stable for some positive value K_1. Therefore θ_x will oscillate and there is a sinusoidal value for e(t) such as $e = E \sin \omega t$. But with any value for E the describing function is a simple gain and we have already shown that for any gain, the system cannot have any sustained oscillation for $K_1 \neq 0$. Consequently, the conjecture is disproved. It can easily be seen that the conjecture is true if $K_1 = 0$.

4 GUIDANCE OF SPACE VEHICLES

The transfer of a vehicle from a point in space and a set of initial values for velocities and accelerations to a desired location satisfying certain conditions on final velocities, etc., is the function of the guidance of the vehicle. It can be accomplished in two ways: a) open-loop, and b) closed-loop.

4.1 Open-Loop Guidance

In this scheme, a trajectory which the vehicle should follow in space is determined a priori. This may be a trajectory satisfying the desired initial and terminal conditions or, in addition, may be an optimal trajectory in some sense. For example, it may accomplish the mission with minimum amount of fuel expenditure, or in minimum time. In cases where the final conditions cannot be achieved exactly, the trajectory may result in minimum value for some function of the terminal miss distance. The trajectory is a function of the mass of the vehicle, maximum value of available torque, initial and terminal conditions and the nature of desired optimality. For example, in case of solar sailing when a

vehicle is leaving Earth orbit and reaches Mars orbit in minimum time, two optimum trajectories result depending on whether the terminal velocities of the vehicle should match that of the orbit or they can assume any arbitrary value. Obviously, when the terminal velocities are to be matched, the optimum time of transfer is longer since this matching is an additional constraint.

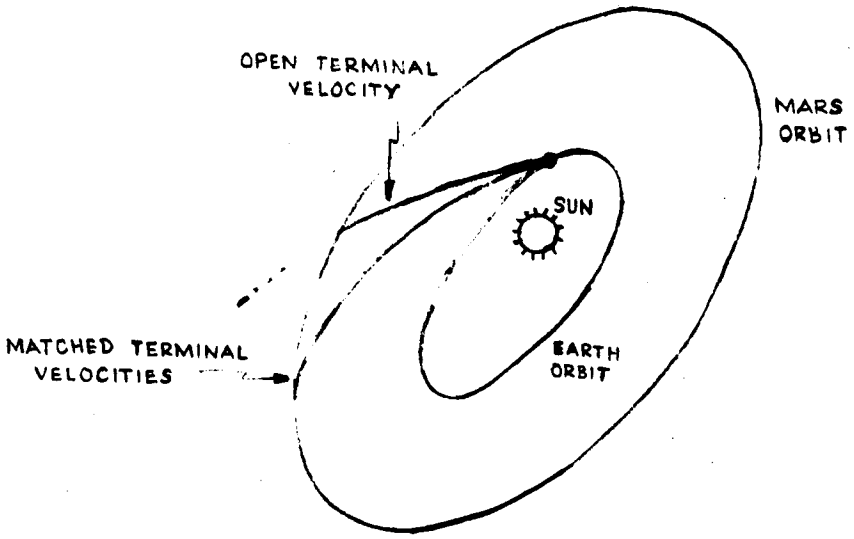


Figure 4.1. Solar Sailing Example

When, by some method, a desired trajectory has been determined, with respect to an inertial coordinate system, in order to maintain the vehicle on that trajectory, it is required to determine the instantaneous position of the vehicle with respect to that coordinate system. This can be accomplished by two means: a) use the rate gyro's and accelerometers on the vehicle to compute the trajectory which the vehicle is following, b) to track the vehicle from another position whose movement with respect to the coordinate system is known.

In case where, at some time, the vehicle trajectory and desired trajectory do not coincide, it is necessary to make trajectory corrections by using the on-board engines. A swiveling engine or body fixed engine can be used to apply thrust (torque) in appropriate direction, the difference being that in the latter case the vehicle orientation is determined by the direction of required correction.

4.2 Closed-Loop Guidance

In many missions, the requirements on the accuracy of achieving the terminal conditions is such that an open-loop guidance is not satisfactory. This

may be because of the movement of the destination point (target) or errors in determining the vehicle trajectory, or difficulties in making exact trajectory correction. Furthermore, it is always desired to make vehicles more autonomous (self governing). Consequently, in many cases, the open loop guidance either is not used at all, or it is only used during the midcourse, that is, transferring the vehicle from the initial conditions to some vicinity of the destination and then performing the last portion of guidance in "closed loop' fashion. A number of examples of closed loop guidance is given in the following paragraphs.

4.2.1 Pursuit

When an intercepting vehicle with initial velocity vector V_I tries to intercept with a target vehicle, with say, a constant velocity V_T, one possibility is that the interceptor orients its velocity vector such that it always goes through the instantaneous position of the target. This is the example of the dog chasing a rabbit. As a two dimensional example, let x(t) be the target path (straight line) with the initial conditions

$$x(0) = x_T(0) \quad , \quad \dot{x}(0) = V_T \quad ; \quad \left\{ x(t) = \begin{bmatrix} x\ (t) \\ \\ x\ (t) \end{bmatrix} \right\}$$

Let the interceptor be at $x_I(0)$ at t = 0. The trajectory that the interceptor will follow is sketched in Figure 4.2.

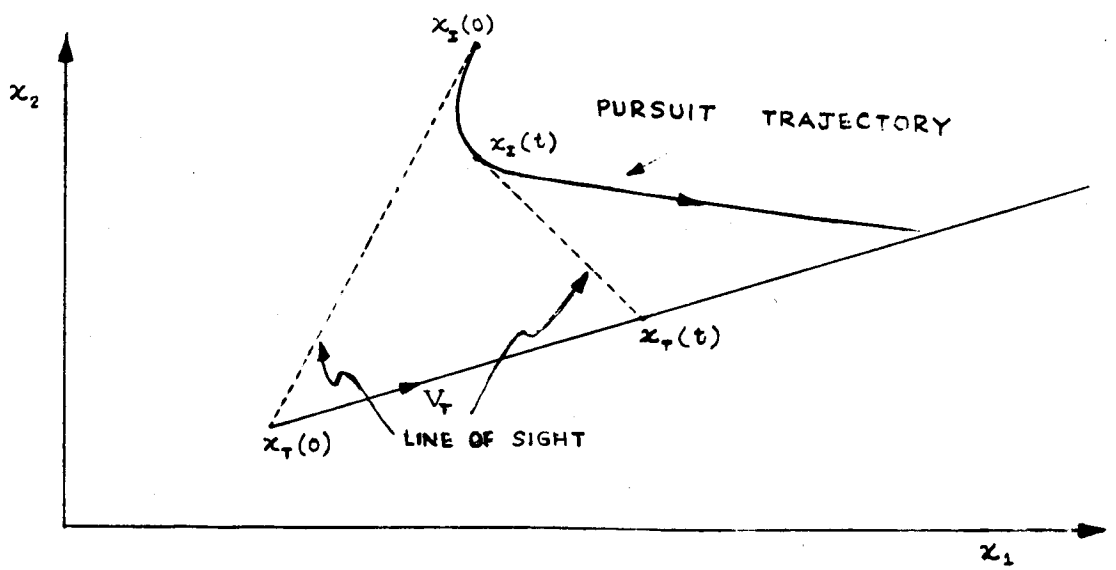


Figure 4.2. Pursuit Navigation

It is clearly seen that, although the target has a very simple flight path, the interceptor has to go through a reasonably complicated maneuver. This

is mainly due to the inefficiency of pursuit navigation. The interceptor has to continuously apply thrust perpendicular to instantaneous velocity vector in order to keep the velocity vector coincident with the line-of-sight (line-of-sight is the vector from instantaneous interceptor position to target position, this is abbreviated by LOS and is given with respect to an inertial coordinate system). A change in direction of the velocity vector can be obtained by applying acceleration in a direction perpendicular to the velocity vector which is accomplished by either orienting the swiveling engine or the whole vehicle in case of body fixed jets in the appropriate direction. A mathematical model for this system is derived in the following. Figure 4.3 represents a schematic diagram of the space vehicle target system. The interceptor should attempt to maintain the orientation of its velocity vector in the appropriate direction. A change in the orientation of V_I is accomplished by applying force perpendicular to the direction of V_I. This force will produce a rate of change of the angle σ, i.e., $\dot{\sigma}$ is proportion to the applied thrust.

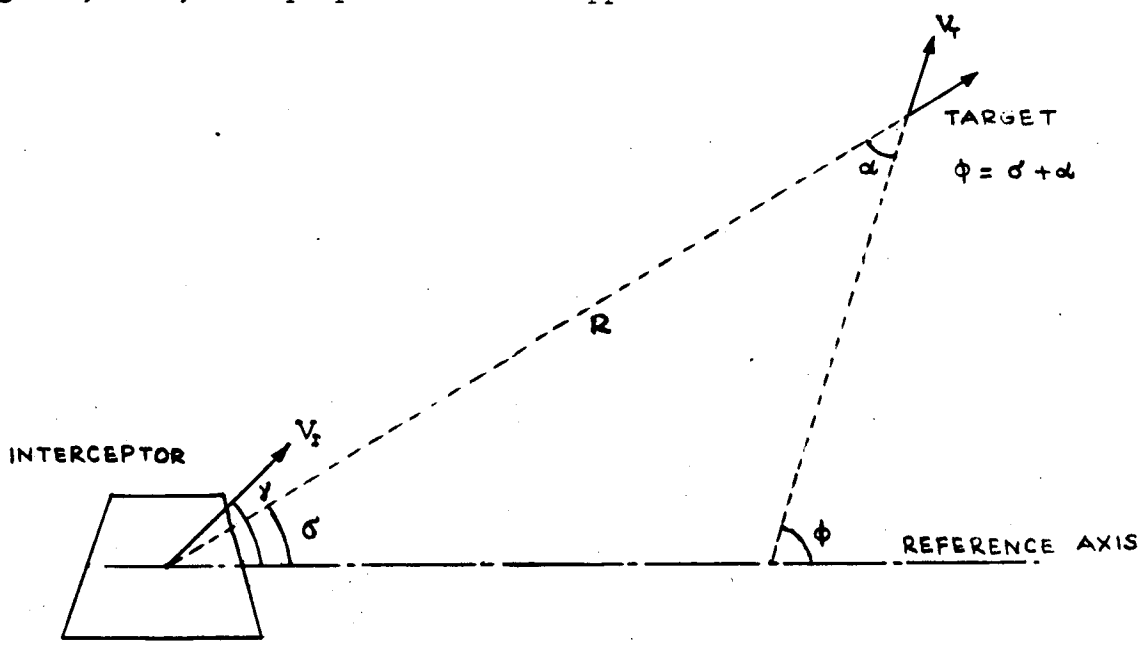


Figure 4.3. Schematic Diagram for Interceptor-Target System

Let

V_I interceptor velocity (constant in magnitude)

V_T target velocity (constant in magnitude)

a_{Tn} the component of target acceleration perpendicular to LOS

R range

Let us first assume $a_{Tn} = 0$. We have the following equations

$$R\dot{\sigma} = V_T \sin(\phi - \sigma) - V_I \sin(\gamma - \sigma) \tag{4.1}$$

$$\dot{R} = V_T \cos(\phi - \sigma) - V_I \cos(\gamma - \sigma) \tag{4.2}$$

Differentiating, 4.1 yields

$$\dot{R}\dot{\sigma} + R\ddot{\sigma} = -\dot{\sigma}V_T \cos(\phi - \sigma) - \dot{\gamma}V_I \cos(\gamma - \sigma) + \dot{\sigma}V_I \cos(\gamma - \sigma) \tag{4.3}$$

Substituting 4.2 into the right hand side of 4.3 results

$$\dot{R}\dot{\sigma} + R\ddot{\sigma} = -\dot{\sigma}\dot{R} - \dot{\gamma}V_{IR} \tag{4.4}$$

where

$$V_{IR} = V_I \cos(\gamma - \sigma)$$

= Component of interceptor velocity in the direction of LOS

From 4.4

$$\ddot{\sigma} = -\frac{2\dot{R}}{R}\dot{\sigma} - \frac{V_{IR}}{R}\dot{\gamma} \tag{4.5}$$

If a_{Tn} was not zero, it would simply add to the $\ddot{\sigma}$ a quantity equal to a_{TN}/R, hence

$$\ddot{\sigma} = -\frac{2\dot{R}}{R}\dot{\sigma} - \frac{V_{IR}}{R}\dot{\gamma} + \frac{a_{Tn}}{R} \tag{4.6}$$

In order to develop a block diagram, let us assume that over small ranges in time, $\dot{R}/R = -V$ = Constant and V_{IR} and R are also constant. Taking the Laplace transform of 4.6 yields ($V_{IR}/R = C$)

$$(s^2 - 2V\,s)\,\sigma = -C\,s\gamma + \mathcal{L}\frac{a_{Tn}}{R} \tag{4.7}$$

This yields the following block diagram

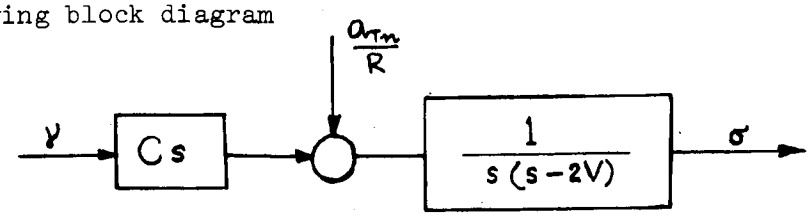


Figure 4.4

The loop is closed through the navigation law which in this case attempts to line up V_I with LOS. This can be accomplished by varying γ proportion to $\gamma - \sigma$ and in a direction which reduces the magnitude of $\gamma - \sigma$. For example

$$\dot{\gamma} = -\lambda(\gamma - \sigma) \tag{4.8}$$

λ is a constant called navigation constant. From (4.8)

$$\gamma(s + \lambda) = \lambda\mathcal{L}\sigma \tag{4.9}$$

Finally the following block diagram is obtained

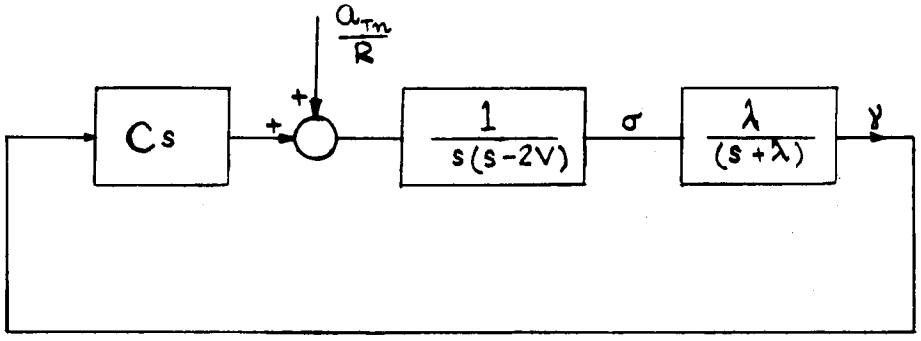


Figure 4.5. Block Diagram for Pursuit Navigation

It is clearly seen that since the loop tries to make $\sigma = \gamma$ and since σ is varying during the flight, consequently the interceptor has to apply acceleration all during the flight.

4.2.2 Proportional Navigation

The proportional navigation is the guidance policy which keeps the LOS non-rotating. Let the LOS angle with respect to some fixed coordinated system be σ and its derivative (LOS rotation) $\omega = \dot{\sigma}$

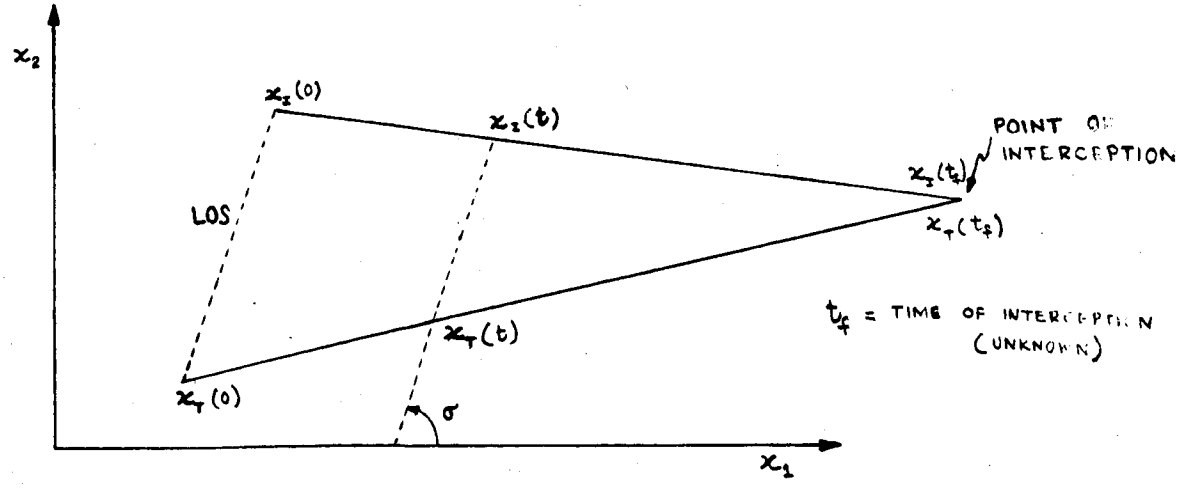


Figure 4.6. Proportional Navigation

From Figure 4.6, it is evident that if LOS is non-rotating, the interception will occur at some time $t_f > 0$. The closed loop system is mechanized such that it "nulls" out any existing LOS rate. The information about the LOS rate is obtained by the tracking system. The equations of motion and a block diagram for the system is derived in the following.

The derivation of Equation (4.6) is applicable here. The difference is only in the navigation law. Here the control loop tries to zero out any rotation of LOS, namely $\dot{\sigma}$. This is done by letting

$$\dot{\gamma} = -\lambda\dot{\sigma} \tag{4.10}$$

Substituting (4.10) in (4.6) yields

$$\ddot{\sigma} = -\frac{2\dot{R}}{R}\dot{\sigma} + \frac{\lambda V_{IR}}{R}\dot{\sigma} + \frac{a_{Tn}}{R} \tag{4.11}$$

Which yields the following block diagram

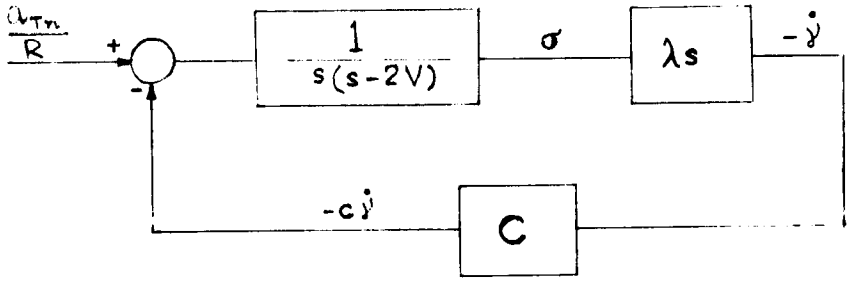


Figure 4.7. Block Diagram for Proportional Navigation

The transfer function between a_{Tn}/R and $\dot{\gamma}$ is

$$\frac{-\dot{\gamma}}{\alpha_{Tn/R}} = \frac{\frac{\lambda s}{s(s-2V)}}{1 + \frac{\lambda C s}{s(s-2V)}} = \frac{\lambda s}{s^2 + s(\lambda C - 2V)} = \frac{\lambda}{s + \lambda C - 2V} \tag{4.12}$$

It is then clear that since it is desired that the required thrust be a diminishing quantity, the system should be stable meaning that

$$\lambda C - 2V = \lambda\frac{V_{IR}}{R} + 2\frac{\dot{R}}{R} > 0 \tag{4.13}$$

or

$$\lambda > \frac{-2\dot{R}/R}{V_{IR/R}} = -\frac{2\dot{R}}{V_{IR}} \tag{4.14}$$

The quantity

$$\Lambda = \frac{\lambda}{-\dot{R}/V_{IR}}$$

is referred to as effective navigation constant and the condition for stability is then $\Lambda>2$. It is clearly seen that when the interceptor is on a collision course and $a_{Tn} = 0$, the quantity $\dot{\gamma}$ remains zero for the entire flight which indicates that no acceleration is required.

5 TRACKING SYSTEMS

In the preceding part, it was pointed out that for closed-loop guidance of vehicles, a measurement of either line-of-sight angle or angle rate with respect to a fixed reference is necessary. This job is accomplished by a tracking system (usually called angle tracking system).

A tracking system is composed of two parts: a) an error detector, which is a device producing a signal proportion to the deviation of antenna center beam from LOS, and b) a control loop which drives the antenna in order to reduce the tracking error to zero. If the tracking error is maintained very small during the flight, the position of the antenna, which is known (can be measured) gives the information about σ or $\dot{\sigma}$ (LOS angle or rate).

5.1 Error Detectors

The following are three important examples of error detectors.

5.1.1 Radar Lobing

Figure 5.1 represents the propagation pattern of a radar antenna.

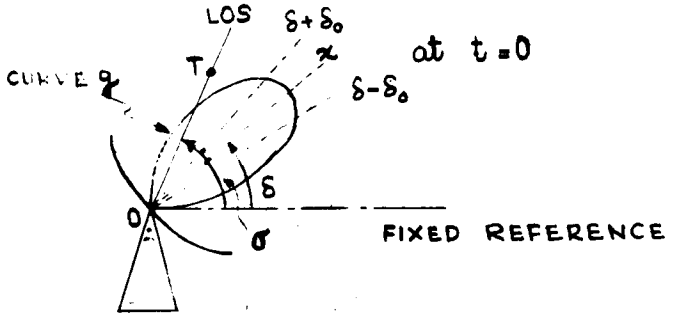


Figure 5.1. Radar Antenna

Since δ is known, the object is to generate a signal proportion to $\sigma - \delta$ where T is the target point. When T is on ox a maximum signal is returned. The farther T is from ox in angle, the smaller will be the amplitude of the return signal. This functional relationship is shown by the curve g in Figure 5.1 and can be represented as

$$z = Kg(\sigma - \delta) \tag{5.1}$$

where z is the radar output and is the analytical representation of the lobe and K is proportion to the length OT. Now let us move the antenna mechanically, or electronically, by the following rule

$$\delta = \delta_0 + \delta_1 \sin wt \tag{5.2}$$

where δ_1 is a constant known quantity. It can be seen that

$$z = - K_1(\sigma - \delta_0) \sin wt \tag{5.3}$$

Multiplying z by sin wt and keeping the d.c. term yields

$$z_{d.c.} = - \frac{1}{2} K_1(\delta - \delta_0) \tag{5.4}$$

This is the required result. Notice that $z_{d.c.}$ contains the sign of $\sigma - \delta_0$, i.e., the position of T with respect to center of the lobe (δ_0).

5.1.2 Error Detection with Fixed Antennas

In certain applications, it is necessary not to mechanically, or electronically, lobe the antenna. In this case two antennas can be used to determine the phase difference between the received signals (Figure 5.2).

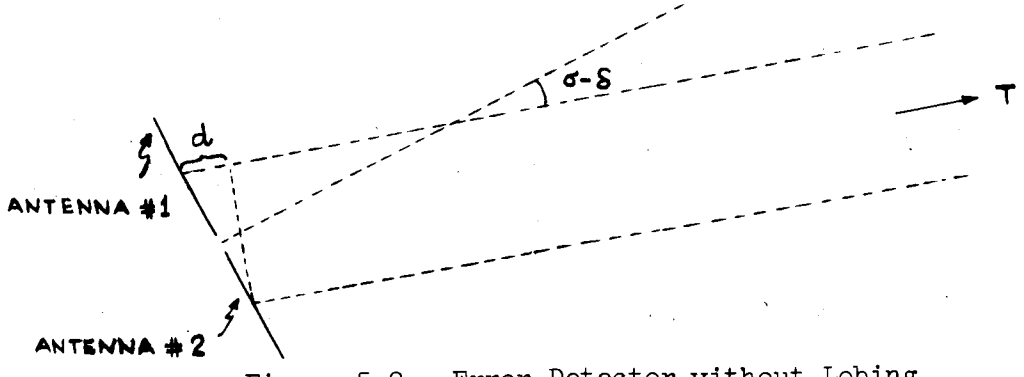


Figure 5.2. Error Detector without Lobing

It is easily seen that the distance d is proportion to $\sigma - \delta$ angle. The incident lines from target are almost parallel lines because the size of combination antennas is much smaller than the distance to target. The two outputs of the antennas are

$$z_1 = K \sin wt \tag{5.5}$$

$$z_2 = K \sin(wt + d) \tag{5.6}$$

Phase shifting z_1 by 90° and then multiplying z_1 and z and retaining the d.c. component yields

$$\text{d.c. component} = \frac{K^2}{2} \sin d \tag{5.7}$$

For small angles $(\sigma - \delta)$ approximately

$$\text{d.c. comp.} = \text{proportional to } \sigma - \delta$$

Again notice that the d.c. component in (5.7) retains the sign information in $\sigma - \delta$.

5.1.3 Optical Error Detector

The idea of lobing can be used in an optical detector. Let four detectors be placed as shown in Figure 5.3.

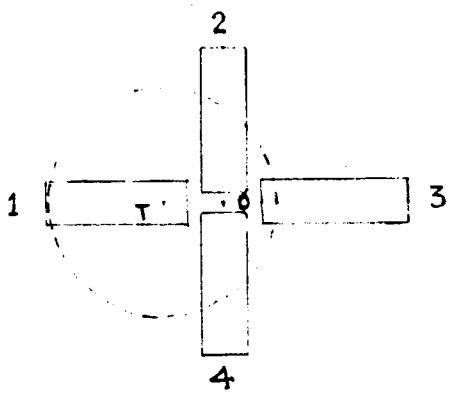


Figure 5.3. Optical Error Detector

Let the target image be at point T without lobing. Lobing therefore will move the image on a circle with center T. When the image is on a detector, an output is produced. The series combination of four outputs of the detectors for the case of Figure 5.3 is presented by Figure 5.4(a). If T coincides with O the center of the detectors, then the case of Figure 5.4(b) will result. Consequently, information about angle of incident light from the target can be obtained from z

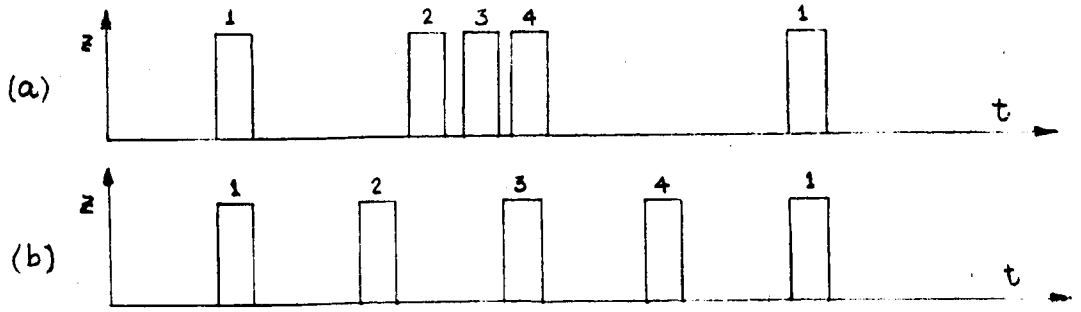


Figure 5.4. Output Waveform of Optical Detector

5.2 Tracking Loop

Any of the above schemes will yield a relationship between the detector output and $\sigma - \delta$ which typically can be represented as in Figure 5.5

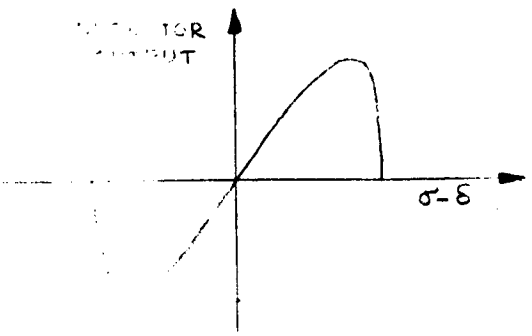


Figure 5.5. Characteristic of Error Detectors

The S-shaped form is because when $\sigma - \delta$ is larger than certain value, the target falls completely outside the radar or optical beam.

Figure 5.6 represents the block diagram of a complete tracking system

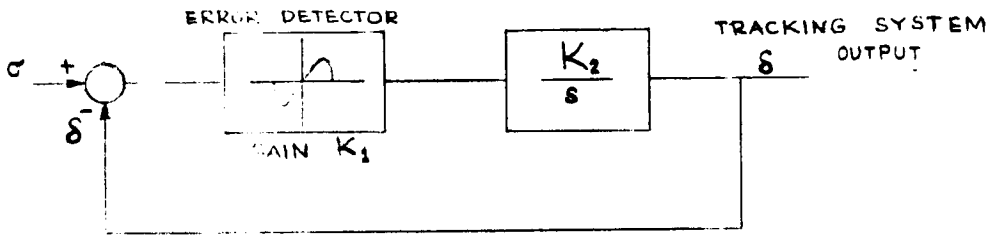


Figure 5.6. Tracking System

over the linear range of the error detector, we have

$$\mathcal{L}\delta = \mathcal{L}\sigma \frac{K_1 K_2/s}{1 + K_1 K_2/s} = \frac{K_1 K_2}{s + K_1 K_2} \mathcal{L}\sigma \qquad (5.8)$$

For example if the antenna is initially not aligned with the target (i.e., $\sigma \neq \delta$) then from 5.8

$$\delta(t) = \sigma\left(1 - e^{-K_1 K_2 t}\right) \qquad (5.9)$$

which shows that as $t \to \infty$, $\delta(t) \to \sigma$, the desired position. In practice, the terms K_2/s should be modified to include the antenna dynamics, and also any filtering of the noise, which may help the performance of the system.

5.3 Frequency Tracking System

Tracking systems other than angle tracking are also in wide use such as the case of tracking the relative speed (range rate) of two vehicles. One

way this can be accomplished is by use of doppler shift discussed below.

If a periodic wave asin wt is reflected back from a stationary incident surface its form will not change. However, if the reflector moves toward or away from the source of transmitted wave then the reflected wave indicates a change in frequency. This frequency change is referred to as Doppler shift.

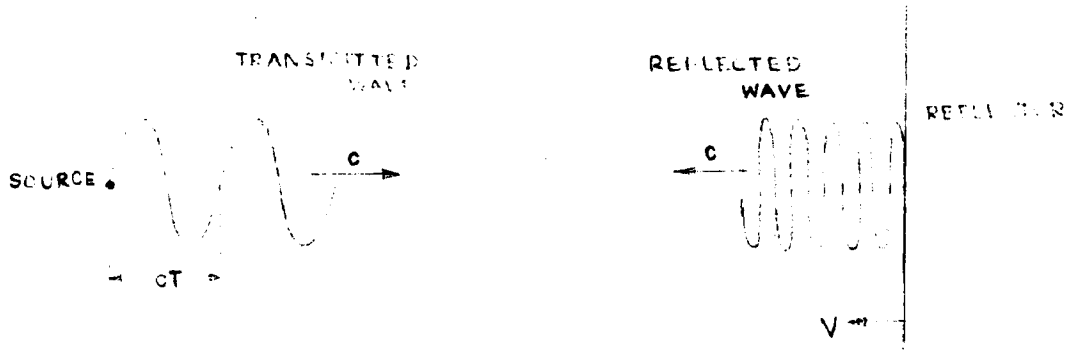


Figure 5.7

If the reflector moves toward the source with a velocity V the period of the reflected wave T' is then

$$cT' = cT - VT \tag{5.10}$$

or

$$f' = \frac{1}{T'} = f(1 + \frac{V}{c}) \tag{5.11}$$

or

$$\Delta f \simeq \text{doppler shift} = \frac{fV}{c} \tag{5.12}$$

It is then evident that if it is desired to generate a signal proportion to V we have to track the frequency f' in the reflected wave of the form

$$b \sin(2\pi f't) \tag{5.13}$$

This is accomplished by the following control loop

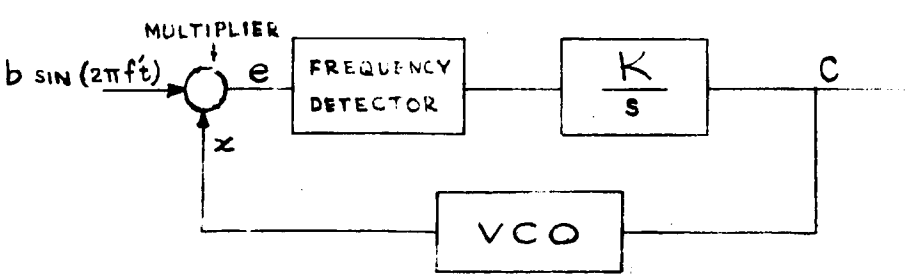


Figure 5.8

VCO (voltage controlled oscillator) is a device with the property that its output is a cosine wave with a frequency proportion to input C

$$x = b_1 \cos 2\pi K_1 C t = b_1 \cos 2\pi f'' t \tag{5.14}$$

Consequently

$$e = b_1 b \sin 2\pi f' t \cos 2\pi f'' t \tag{5.15}$$

$$= \frac{b_1 b}{2} \left\{ \sin 2\pi (f' - f'') t + \sin 2\pi (f' + f'') t \right\} \tag{5.16}$$

The block diagram can then be modified

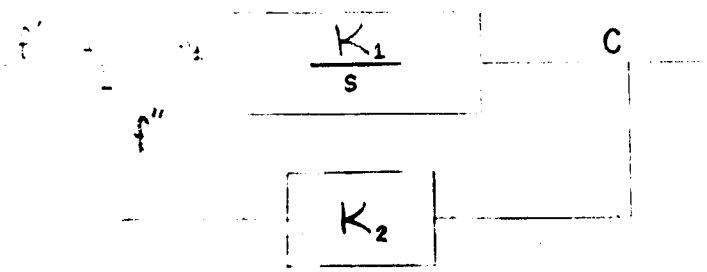


Figure 5.9

The system is stable and

$$\lim_{t \to \infty} e_1(t) = 0$$

which means after the transients have died out

$$C = \frac{1}{K_2} f' \tag{5.17}$$

and is the desired result.

6 STUDY OF A CONTROL SYSTEM

6.1 Surveyor Moon Landing Control System

In this section a rather sophisticated system is chosen for study. In brief the various phases of the vehicle motion are the following.

After launch the vehicle attitude is locked to Sun-Canopius reference system. By means of the on board attitude control system the vehicle attitude is adjusted at the time of midcourse correction. After the appropriate correction is made the attitude is locked to Sun-Canopius reference again. This continues up to the vicinity of the moon. Retro rockets are used to slow down the vehicle (again the vehicle attitude is maintained in appropriate direction for the retroaction). At 40 feet from surface of the moon the attitude is held con-

stant and the longitudinal velocity is controlled to reach 5 ft/sec and is maintained until 14 feet from the moon surface. At this time all engines are turned off and the vehicle is let to drop. The touch down velocity is about 12.5 ft/sec.

Figure 6.1 represents the phase summary of the vehicle flight. Figure 6.2 is the block diagram for the longitudinal velocity control system. Figure 6.3 is the block diagram of the pitch and yaw attitude control systems. Figures 6.4 and 6.5 are the corresponding analog computer simulation diagrams.

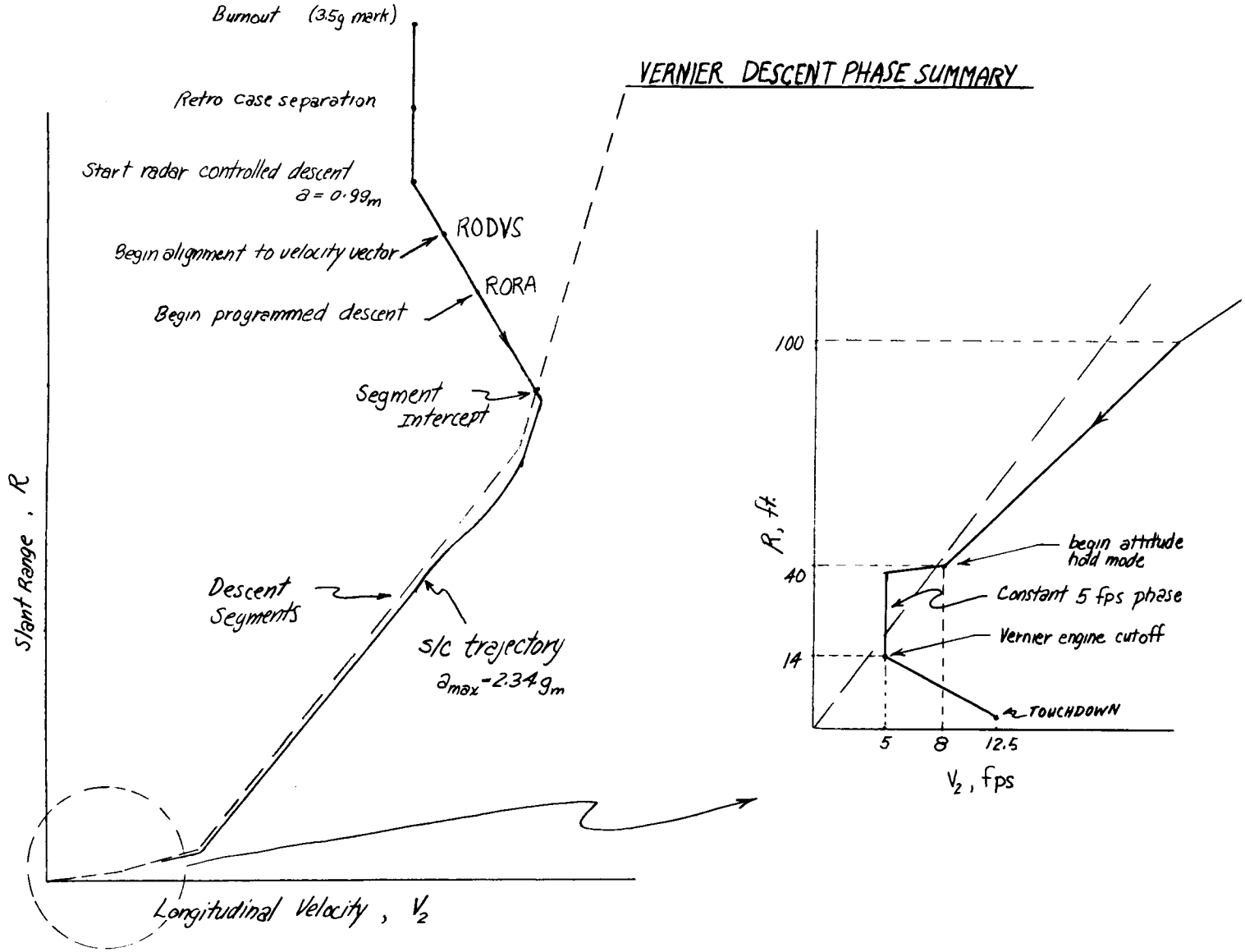

Burnout (3.5g mark)
Retro case separation
Start radar controlled descent
a = 0.99m
RODVS
Begin alignment to velocity vector
RORA
Begin programmed descent
Segment intercept
Descent Segments
s/c trajectory
amax-2.34gm
Slant Range, R
Longitudinal Velocity, V2
VERNIER DESCENT PHASE SUMMARY
100
40
14
R, ft.
begin attitude hold mode
Constant 5 fps phase
Vernier engine cutoff
TOUCHDOWN
5
8
12.5
V2, fps

Figure 6.1

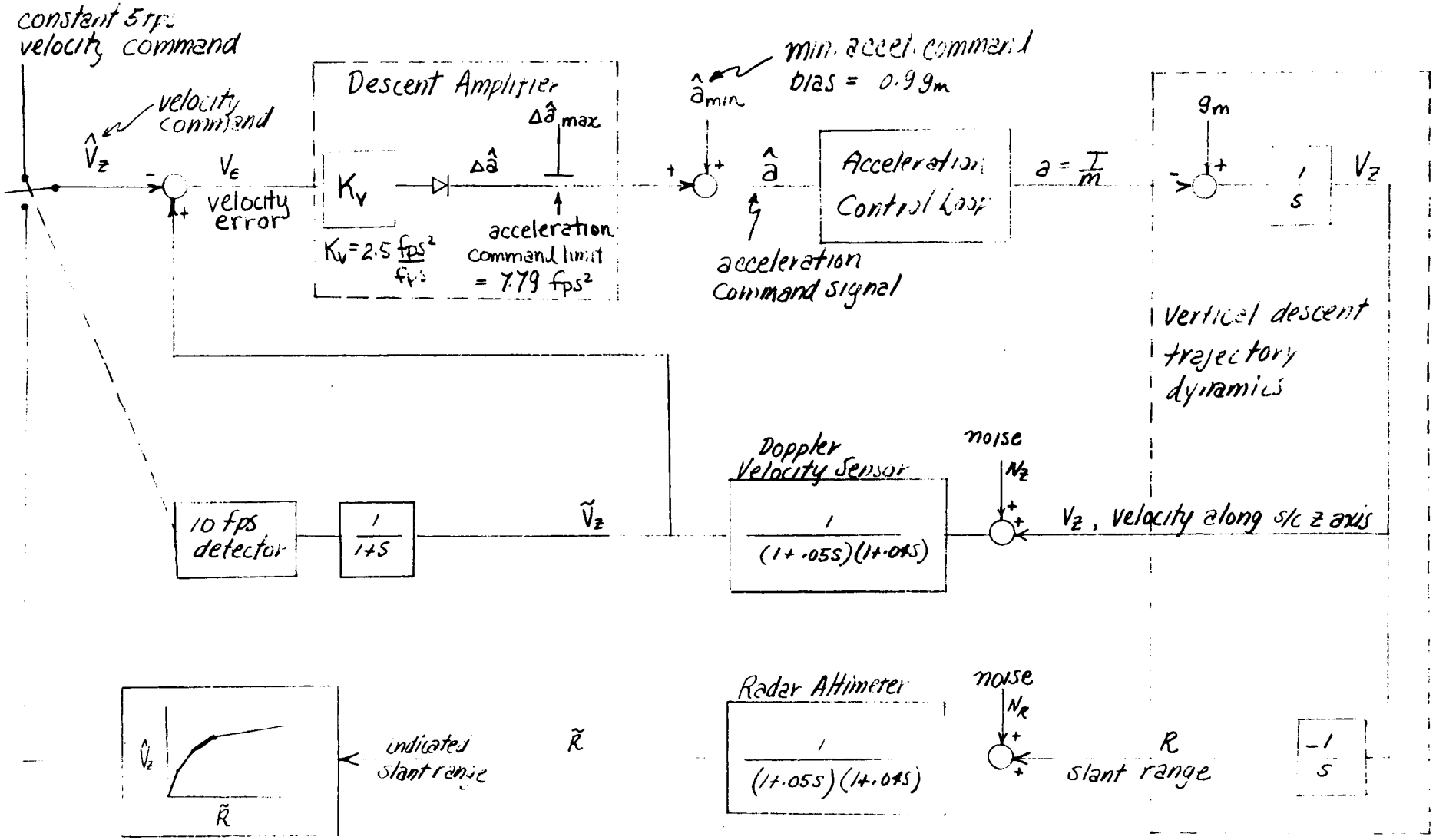


Figure 6.2. Longitudinal Velocity Control System

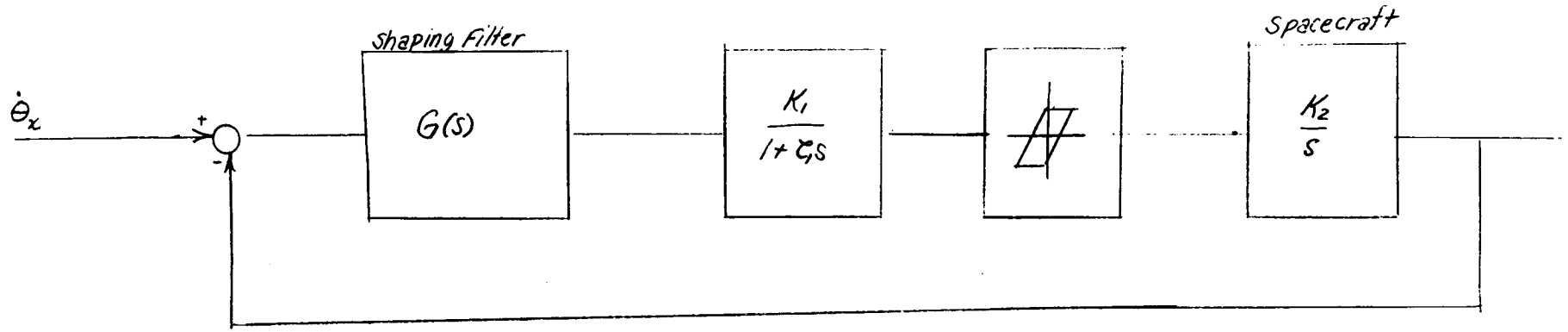


Figure 6.3. Pitch and Yaw Attitude Control

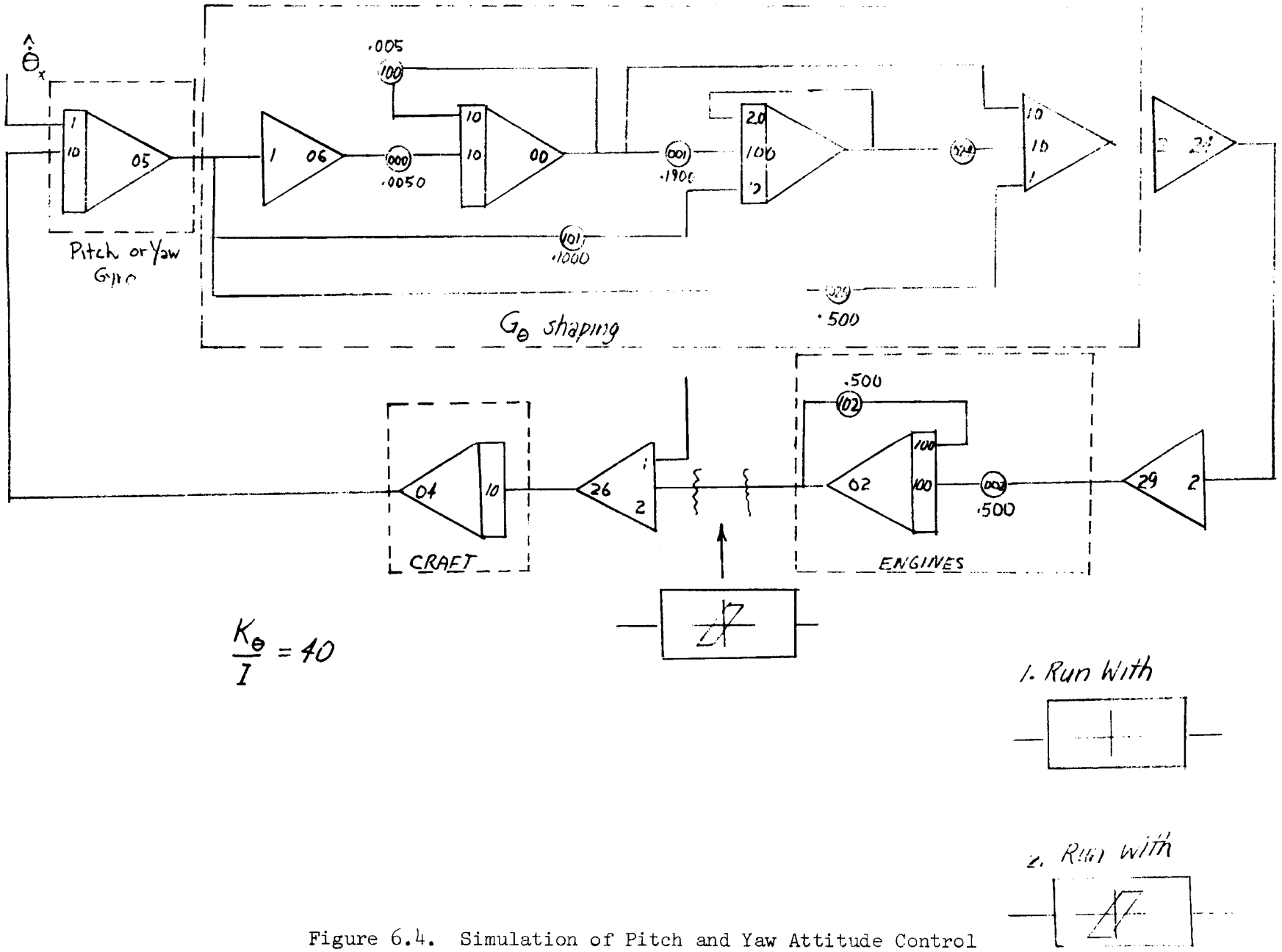
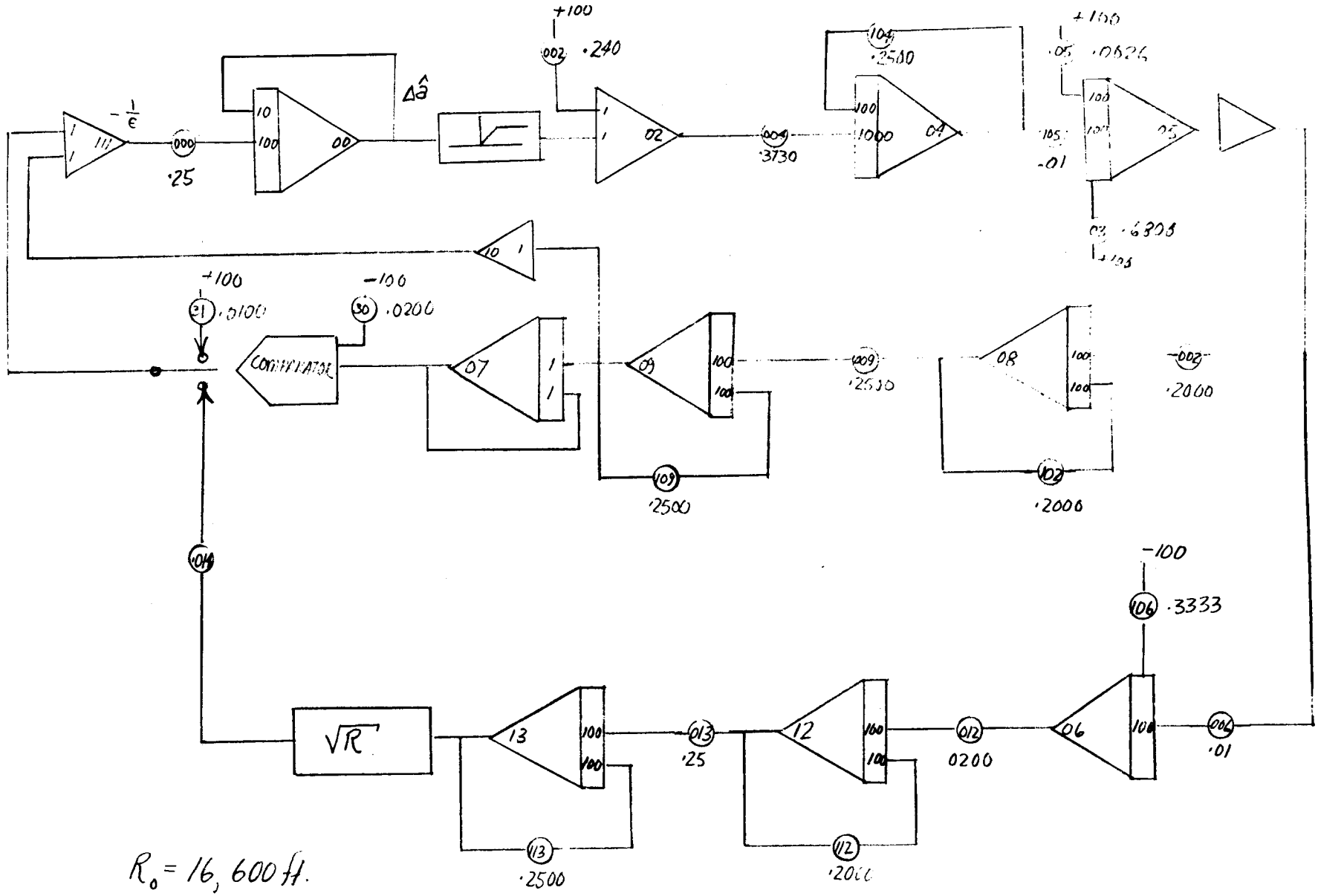


Figure 6.4. Simulation of Pitch and Yaw Attitude Control



$R_o = 16,600\ \Omega$.

$V_o =$

Figure 6.5

III SYSTEM ENGINEERING (C. CASANI)

1 INTRODUCTION TO SYSTEM ENGINEERING

Although this chapter will be oriented primarily toward System Engineering as used at the Jet Propulsion Laboratory to accomplish its space flight missions, classical System Engineering will also be considered. This material will provide the student with a better background and insight into the type or classes of problems that can be encountered in the field.

This course will be divided into two parts: an introduction to system engineering and development of some of the classic tools of system engineering such as linear programming, game theory, probability theory, decision theory and logic; and some of the practical considerations of system engineering as used at the Jet Propulsion Laboratory. This will include in particular the design of a hypothetical 1971 Mars Mission.

In the two decades since World War II, the field of System Engineering has grown immensely. The war represented a large and complex operating system which was so difficult to understand that decisions could not be made on intuition only, but rather required rigorous analysis; thus, operations research came into being, and tools like linear programming found practical applications.

One of the by-products of the war was the missile and out of that has evolved the present day Aerospace Industry, which has been instrumental in the refinement of system engineering. Today, system engineering has become not only a widely accepted term throughout the industry, but also an accepted field of major in many colleges and universities.

Although the term is widely used and accepted, it is often misused and misunderstood. The terms, system engineering, system analysis, operations research and operation analysis are often confused and used interchangeably. Succeeding paragraphs will define these terms as they will be used in this course. These definitions may not be universal, but they are what we will use and therefore in our world they are correct. Since our world is the only correct world, these then are the only acceptable definitions.

a. Operations Research

Operations Research is the development of such mathematical tools as linear programming, game theory, queueing theory, and decision theory, all of which are used throughout the field. The development of these tools is usually done on a mathematical level or at the research level. Their application is usually more classical than practical.

b. Operations Analysis

Operations Analysis is the development and analysis of mathematical models of an operating system. The Military Services and large corporations such as the airlines, the oil industry, and the shipping industry use these skills to improve the operating efficiency of the organization. Some of the classis mathematical tools used in operations research are used there also.

c. System Engineering

System Engineering is the conception, design, development, and operation of large complex systems. It is concerned with creating an optimum system to perform a specific job within a set of specific constraints, the most important of which are usually performance, time, cost, and manpower. It is this concept of System Engineering which is most familiar throughout the Aerospace Industry.

d. System Analysis

System Analysis is a discipline which models a system by describing the inter-relationships between its various elements, and tradeoff studies are an important part of this discipline. The development of the required software of the system is part of System Analysis.

At this point, it is important to have a good understanding of exactly what the word "system" encompasses. Figure 1.1 is a graphic representation of any system: an operating system whose input is the performance requirements, and whose output is the actual performance and reliability.

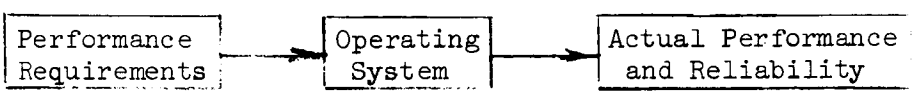


Figure 1.1

An operating system is composed of the following elements:

1) Hardware
2) Operating Procedures
3) Personnel and Training
4) Environment
5) Support Equipment and Logistics
6) Cost

In discussions of these types of systems, the word, "optimization," is often used, and it is important to understand this term also. There are two types of optimizing:

1) Fixed performance - minimum cost

2) Fixed cost - maximum performance

In both types, performance includes reliability. In the first case, we fix the performance and then the optimum solution is one which meets the performance at the minimum cost. In the second case, we fix the cost and then the optimum solution is one which maximizes the performance.

1.1 System Engineering Process

Let us consider the actual process of system engineering. As defined earlier, system engineering includes the conception, design, development and operation of a system. This process consists of the following major phases:

a. Project Definition

b. System Design

c. Preliminary Design

d. Hard Design

e. Fabrication

f. Assembly and Test

g. Operation

1.1.1 Project Definition Phase

Prior to project definition, the general requirements and general feasibility of this project must be demonstrated. The method used for this demonstration is usually rather unorthodox, and the logic is often difficult to understand. There have been many projects proposed and undertaken ostensibly to fulfill some specific military or civic need, which were actually only to fill someone's pocket or ballot box. Be that as it may, the problem we want to look at is given a set of project requirements. How does one go about accomplishing the project? These requirements are often stated in very general terms: "Take close-up pictures of Mars in 1964" or "Put a man on the Moon before 1970." It is often the case that the bigger the assignment, the more general the requirements.

During the project definition phase, the constraints and ground rules under which the project is going to operate are defined: the management tools, the overall schedule, the manpower and the budget. A set of project objectives must be developed and agreed upon. These objectives should be clearly stated

and understood by the project management and the customer. This is a most important step, since it is against these stated objectives that the project will be continually reviewed and finally appraised. The project definition phase is usually extremely short in comparison to the life of the project, yet some of the most critical decisions which shape the entire project must be made during this phase.

1.1.2 System Design Phase

The system design phase consists of defining a system capable of fulfilling the project objectives. This system should be defined in total with particular emphasis on those areas in which extensive development is anticipated. A system level of understanding is required throughout all elements of the system. This is not to imply that the detail mechanization of the subsystems are specified, but that their functional requirements are described. At the end of this phase, there should not be any feasibility questions unanswered. If there are, then it is foolhardy to proceed with the project for, when we are concerned with large complex systems, they are very serious in nature, and one missing major element is catastrophic on the final outcome. This activity is sometimes called the Conceptual Design Phase.

1.1.3 Preliminary Design Phase

The Preliminary Design should always begin with a critical design review of the system design. It may be necessary to modify the system design after such a review. During this phase, the subsystems and their interface characteristics are functionally specified. This is an extremely critical phase in the development of a system; to assure the best overall system design, continual tradeoffs between the subsystems must be performed. A proper balance of risks is best achieved by prudent tradeoffs between the subsystems conducted at the system level. In addition, care must be taken to avoid strong intra-subsystem dependence. The system at this level of design should have as much compliance as practical.

1.1.4 Hard Design Phase

During this phase, the detail design of each subsystem is implemented. These designs are developed in accordance with the functional specifications developed in the Preliminary Design. It is important that all subsystems enter this phase of the design at an equal level of design (i.e., good functional specifications), for if one subsystem is poorly understood, its design may be dictated by the others and this may present extreme difficulties in mechaniza-

tion which could have been avoided by some proper tradeoffs earlier in the Preliminary Design.

1.1.5 Fabrication Phase

This phase is the actual manufacturing and inspection of the hardware. The hardware dependent software is developed and checked out during this phase.

1.1.6 Assembly and Test Phase

After fabrication, the subsystems are tested at the subsystem level and then assembled into an entire system. System tests are then conducted to demonstrate the system's capability to meet its performance requirements. Specific tests are conducted to disclose any anomalies at the system level.

1.1.7 Operation Phase

After the system is successfully working as a unit, it is then committed to operations. Systems which have long operational lifetime and are a production line item, such as large missile systems, must have a well-established feed back loop into the design as operational problems are disclosed.

2 LINEAR PROGRAMMING

Linear Programming had its beginnings in World War II and has continued to develop since then. It is a technique which deals with problems that are characterized by being linear in nature and usually having no one unique solution. In other words, the problem can be bound by a set of linear constraints and within this set of constraints, a specific function is to be optimized. Problems of this type usually contain many variables.

The general linear programming problem can be reduced to the following formulation:

To find $x_u(\geq 0)$, $j = 1, 2, 2 \cdots, n$.

Subject to the following constraint:

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, 3, \cdots m \tag{2.1}$$

where $m<n$

such that

$$C = \sum_{j=1}^{n} c_m x_j \text{ is a minimum (or maximum)} \tag{2.2}$$

and where a_{ij}, and b_i and c_j are all given constants.

The equations which are of the form

$$\sum_{j=1}^{n} a_{ij} x_n = b_1 \tag{2.3}$$

are called constraint equations. The equation of the form

$$C = \sum_{j=1}^{n} c_j x_j$$

is called the cost function or sometimes the objective function. It is the value of this cost function we wish to optimize.

The expanded form of the constraint equations is:

$$\begin{aligned}
&a_{11} x_1 + a_{12} x_2 + a_{13} x_3 + a_{14} x_4 + \cdots + a_{1n} x_n = b_1 \\
&a_{21} x_1 + a_{22} x_2 + a_{23} x_3 + a_{24} x_4 + \cdots + a_{2n} x_n = b_2 \\
&a_{31} x_1 + a_{32} x_2 + a_{33} x_3 + a_{34} x_4 + \cdots + a_{3n} x_n = b_3 \\
&\quad\vdots \\
&a_{m1} x_1 + a_{m2} x_2 + a_{m3} x_3 + a_{m4} x_4 + \cdots a_{mn} x_n = b_m
\end{aligned} \tag{2.4}$$

and the cost function:

$$c_1 x_1 + c_2 x_2 + c_3 x_3 + c_4 x_4 + \cdots + c_n x_n = C$$

It can be seen that if m = n, there are m equations and n unknowns. This is, then, a simple problem of m simultaneous linear equations whose solution is straightforward. Now if m>n, that is there are more equations than unknowns, the problem is overconstrained and has no meaningful solution.

The case of interest is where m<n, that is, there are more unknowns that equations. In such a situation, we can arbitrarily choose any values for n - m of the variables. This, then, reduces the problem to one which can be solved by the use of simultaneous equations. Then the trick is to choose these n - m variables such that when the values of all the x_j's are obtained, they give a minimum (or maximum) solution to the cost function.

There are many techniques for solving these types of problems, possibly the most frequently used is the Simplex Method. This method allows rigorous solutions to large linear programming problems, and smaller problems can be solved by inspection and a few rules of thumb.

2.1 The Assignment Problem

Consider a large corporation with branch offices throughout the country. This corporation has four executive job openings: one in New York, Los Angeles, Philadelphia, and Detroit. The corporation has been studying the records of its junior executives and has chosen four men to fill the positions. These men are located in branch offices scattered across the country: one in San Francisco, Chicago, New Orleans, and Milwaukee. Now all jobs are equal, and all men are equally suited to fill any job. The problem is to determine the minimum cost of relocating the four men. Moving any one man from his present location to any one of the four job locations has associated with it a fixed cost. It is, therefore, possible to determine the cost of the sixteen possible moves. If we let the jobs be represented by J_j, $j = 1, 2, 3, 4$ and the men by M_i, $i = 1, 2, 3, 4$, then the problem can be represented in the matrix shown in Figure 2.1. The cost numbers are shown in thousands of dollars.

	J_1	J_2	J_3	J_4
M_1	14	5	5	5
M_2	2	12	6	7
M_3	7	8	3	9
M_4	2	4	6	10

Figure 2.1

By inspection, we can see that this problem is of the general form of a linear programming problem, with an additional consideration, we can only send one man to one job.

Now let c_{ij} be the cost index such that c_{ij} is the cost of sending M_i to J_j, and let x_{ij} be the assignment index such that

$$x_{ij} = 1 \text{ if } M_i \text{ is assigned to } J_j \text{ and,}$$

$$x_{ij} = 0 \text{ if } M_i \text{ is not assigned to } J_j$$

It must also be noted that

$$\sum_{j=1}^{4} x_{ij} = 1, \quad i = 1, 2, 3, 4 \tag{2.5}$$

$$\sum_{i=1}^{4} x_{ij} = 1, \; j = 1, 2, 3, 4 \qquad (2.6)$$

These two statements imply that one man can be assigned to one job only; that is, x_{ij} can take the value one only once in each row and column. Then the cost function has the form:

$$C = \sum_{i=1}^{4} \sum_{j=1}^{4} c_{ij} x_{ij} \qquad (2.7)$$

and this double summation, subject to the previous constraints, is to be minimized. In this particular problem, there are 16 unknowns, 7 dependent variables and 9 independent variables.

This problem can be solved by inspection and use of two general principles:

a. The Principle of Least Choice.

b. The Principle of Interference.

Since the cost function is to be minimized, the first obvious choice would be to pick the smallest cost element of the matrix. In column one, the c_{21} and c_{41} elements are both 2. As soon as one element is chosen, by the principle of interference, all other elements of its row and column are eliminated. If the c_{21} element is chosen, then the remaining elements of the second row are eliminated. This appears to be a prudent choice because of the 12 in that row. The next smallest element of this matrix is c_{33}, the 3, and then c_{42}, the 4, and then c_{14}, the 5. Choosing each one of these elements yields the solution shown in Figure 2.2.

	J_1	J_2	J_3	J_4
M_1	14	5	5	5
M_2	2	12	6	7
M_3	7	8	3	9
M_4	2	4	6	10

Figure 2.2

The value of the cost function for this solution is:

$$C = 2 + 4 + 3 + 5 \qquad (2.8)$$

$$C = 14$$

and this is the minimum. To prove that this solution is actually a minimum, arrange all the c_{ij}'s in ascending order.

$$2, 2, 3, 4, 5, 5, 5, 6, 6, 7, 7, 8, 9, 10, 12, 14$$

The first c_{ij} is 2, but this is not independent of the second c_{ij} and therefore only one can be chosen. This dependence is because they are both in the first column. The next four c_{ij}'s are 2, 3, 4, 5, and their sum is 14, the value of C. Therefore, C = 14 must be the minimum solution.

At this point, a short discussion on the principal of interference is proper. These types of problems are of a sequential decision class. That is, for each decision which is made, there are a number of other choices (potential decisions) which are eliminated. It also turns out that the first decision eliminates the most choices, and each subsequent decision a smaller number. This is illustrated in the problem. The original matrix was a 4×4 matrix with 16 possible choices; when c_{21} was chosen, the matrix was then reduced to a 3×3 matrix with only 9 possible choices. The number of available choices is reduced quadrically with each decision. A point to be noted here is the importance of the early decisions.

This method of solution can be used on many such transportation problems. As the size of the matrix increases, the ease with which this method can be used decreases. One such technique is by the use of an established algorithm.

With several fundamental assumptions, which it is not within the scope of this text to prove, we will develop a technique for solving larger assignment matrices. First, we will assume that if a matrix has a set (or sets) of independent elements whose sum is a minimum, that by adding or subtracting a constant from every element in any row (rows) or column (columns), we will generate a new matrix whose minimum is contained in the same original set of independent elements. A set of independent elements is one in which each element is contained in one and only row and column of the matrix.

Secondly, we will assume that by proper manipulation of the matrix, we can develop a set of independent zeros in the matrix. Finally, this set of independent zero elements corresponds uniquely to a minimum solution of the original matrix.

Consider the matrix of the original problem, shown in Figure 2.3.

	J_1	J_2	J_3	J_4
M_1	14	5	5	5
M_2	2	12	6	7
M_3	7	8	3	9
M_4	2	4	6	10

Figure 2.3

Now, subtracting the smallest element from each column, a new matrix is developed (Figure 2.4). This matrix contains a set of independent zeros which do correspond to the original solution as shown in Figure 2.2.

	J_1	J_2	J_3	J_4
M_1	12	1	2	0*
M_2	0*	8	3	2
M_3	5	4	0*	4
M_4	0	0*	3	5

Figure 2.4

This serves to illustrate the general idea, but with a rather simple matrix. Figure 2.5 shows another matrix in which the answer will not fall out quite so easily.

2	6	5	9
3	4	8	8
5	1	2	3
4	3	2	7

Figure 2.5

First, subtract from each row its smallest element (see Figure 2.6).

0	4	3	7
0	1	5	5
4	0	1	2
2	1	0	7

Figure 2.6

This does not give zeros in all columns, so subtract from each column its smallest element (see Figure 2.7).

0	4	3	5
0	1	5	3
4	0	1	0
2	1	0	5

Figure 2.7

There now exists at least one zero in each row and column, but only three of them are independent. Now to create another independent zero, add 1 to each element in the 3rd row (see Figure 2.8) and then subtract

0	4	3	5
0	1	5	3
5	1	2	1
2	1	0	5

Figure 2.8

from each column its smallest element (see Figure 2.9). This matrix now contains

0*	3	3	4
0	0*	5	2
5	0	2	0*
2	0	0*	4

Figure 2.9

four independent zeros. These zeros now correspond to the elements of the original matrix whose sum is a minimum. The solution then is

$$C = 2 + 4 + 2 + 3$$

$$C = 11$$

Let us now look at the application to this problem of the algorithm shown in Figure 2.10. This algorithm can be stated in the following seven steps, and when these steps are followed properly, the solution will be achieved.

1. Subtract from each row its smallest element. Subtract from each column its smallest element. Choose a trial set of independent zeros and star them. Go to step 2.
2. Cover each column which has a starred zero. If all columns are covered, the solution is complete; if not, go to step 3.
3. Look for an uncovered zero. If there is none, go to step 7. If one is found, go to step 4.
4. Prime this zero. Look for a starred zero in the same row. If there is one, go to step 5. If there is none, go to step 6.
5. Cover the row, uncover the column of the starred zero, and go to step 3.
6. There now exists a unique chain, starting at the primed zero, going vertically to a starred zero, horizontally to a primed zero, etc., and ending on a primed zero (with no starred zero in its column.) Go through this chain, changing primes to stars and erasing stars. Now erase all primes, uncover all rows, and go to step 2.
7. Find the smallest uncovered element in the matrix. Add this element to the covered rows and subtract it from the uncovered columns (or add it to the covered columns and subtract it from the uncovered rows) (or add it to the twice-covered elements

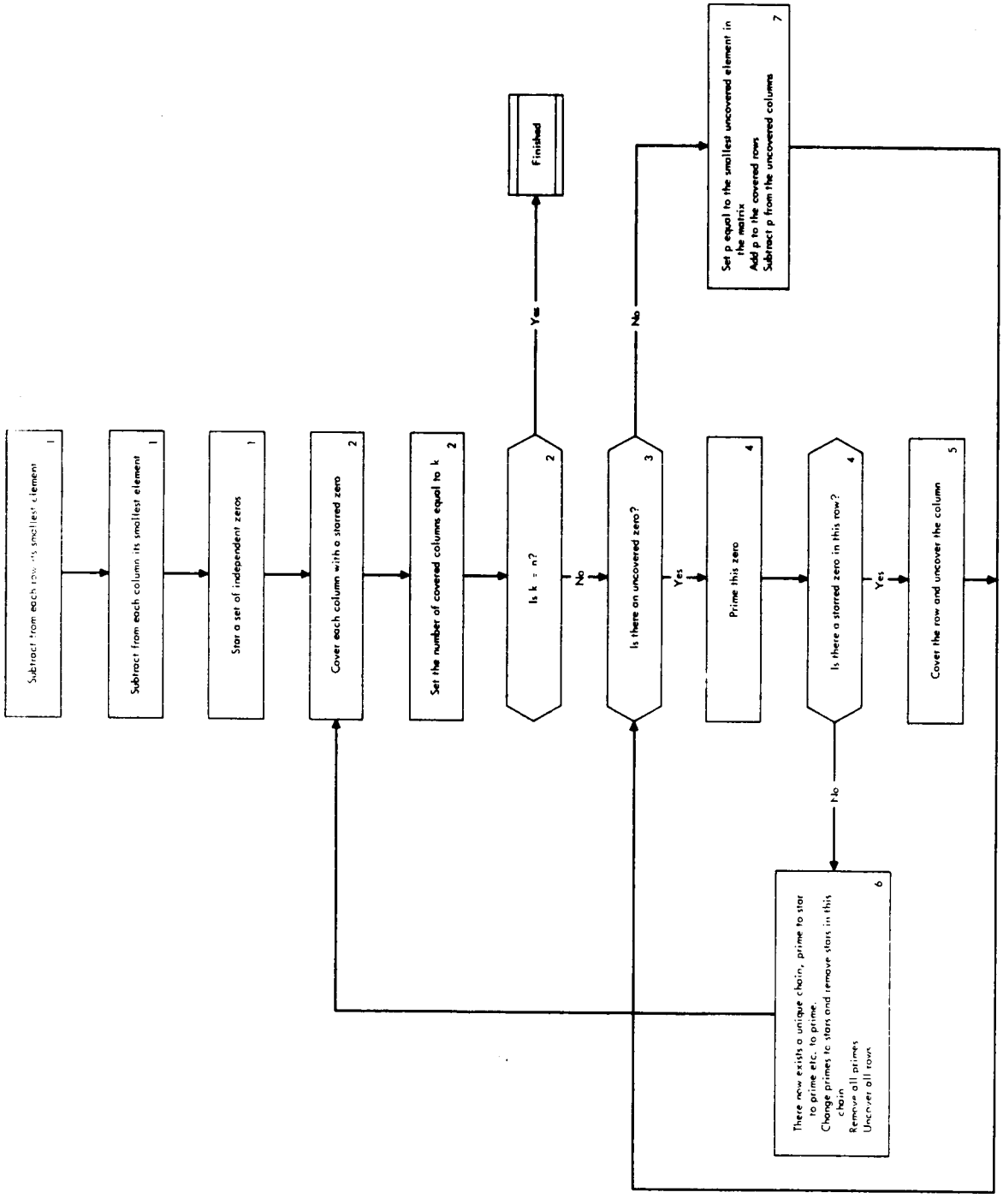


Figure 2.10. Algorithm for Assignment Problem

and subtract it from the uncovered elements). Do not change any stars, primes, or coverings. Go to step 3.

The original problem is shown in Figure 2.11.

2	6	5	9
3	4	8	8
5	1	2	3
4	3	2	7

Figure 2.11

Subtract from each row its smallest element (Figure 2.12).

0	4	3	7
0	1	5	5
4	0	1	2
2	1	0	7

Figure 2.12

Subtract from each column its smallest element (Figure 2.13).

0	4	3	5
0	1	5	3
4	0	1	0
2	1	0	5

Figure 2.13

Choose a trial set of independent zeros and star them (Figure 2.14.

0*	4	3	5
0	1	5	3
4	0*	1	0
2	1	0*	3

Figure 2.14

Cover each column which has a starred zero (Figure 2.15).

x	x	x	
0*	4	3	5
0	1	5	3
4	0*	1	0
2	1	0*	3

Figure 2.15

Look for an uncovered zero. One exists in the fourth column. Prime this zero. A starred zero exists in the same row. Now cover the row and uncover the column of the starred zero (Figure 2.16).

	X		X	
	0*	4	3	5
	0	1	5	3
X	4	0*	1	0'
	2	1	0*	3

Figure 2.16

Now all zeros are covered, step seven now requires us to find the smallest uncovered element of the matrix. This is 1. Add this element to the covered rows (Figure 2.17), and

	X		X	
	0*	4	3	5
	0	1	5	3
X	5	1*	2	1'
	2	1	0*	3

Figure 2.17

subtract it from the uncovered column (Figure 2.18).

	X		X	
	0*	3	3	4
	0	0	5	2
X	5	0*	2	0'
	2	0	0*	2

Figure 2.18

Look for an uncovered zero. One exists in the second column, second row. Prime this zero (Figure 2.19).

	X		X	
	0*	3	3	4
	0	0'	5	2
X	5	0*	2	0'
	2	0	0*	2

Figure 2.19

There are no starred zeros in the second row. There now exists a unique chain, starting at the primed zero, going vertically to a starred zero, horizontaly to a prime zero. Go through this chain, changing primes to stars and removing stars (Figure 2.20).

0*	3	3	4
0	0*	5	2
5	0	2	0*
2	0	0*	2

Figure 2.20

This, then (Figure 2.20) is the solution to the problem since there are four independent zeros. The minimum solution from the original matrix is then

$$C = 2 + 4 + 2 + 3$$

$$C = 11$$

2.2 The Transportation Problem

The transportation problem is another classic problem of linear programming and is similar to the assignment problem in many ways. To get an insight into the problem, let us consider the following situation:

A large bicycle manufacturer has three factories located in three different spots across the country, and at five other locations, he has his warehouses. Now, the cost of shipping one bike from any factory to any warehouse is unequally determined. The problem with which we wish to concern ourselves is the minimum cost of shipping all the bikes from the factories to the warehouses, where it is assumed that each warehouse has a specific demand the number of bikes wanted, and each factory has a specific supply, the number of bikes available. To simplify the problem, let us consider only the case where the total supply is exactly equal to the total demand.

Let us assign some specific values to the problem and develop a solution. Here the problem is represented in a matrix form.

	r_j / a_i	W_1	W_2	W_3	W_4	W_5
		40	40	80	80	80
F_1	100	2	5	2	3	3
F_2	100	2	2	2	1	0
F_3	120	3	6	2	1	4

The a_i's are the amounts available; the r_j's are the amounts required; the elements are the cost of shipping from each factory to each warehouse per unit. Then, in general, we have:

F_i is the i^{th} factory,

a_i is the amount available at the i^{th} factory,

W_j is the j^{th} warehouse,

r_j is the amount required at the j^{th} warehouse,

C_{ij} is the unit cost of shipping from F_i to W_j,

x_{ij} is the number of units shipped from F_i to W_j,

$$\sum_{i=1}^{3} a_i = \sum_{j=1}^{5} r_j, \tag{2.9}$$

$$\sum_{j=1}^{5} x_{ij} = a_i,\ i = 1, 2, 3, \text{ and} \tag{2.10}$$

$$\sum_{i=1}^{3} x_{ij} = r_j,\ j = 1, 2, 3, 4, 5. \tag{2.11}$$

The problem is now to minimize the total cost of shipping all the units. That is to minimize the following:

$$C = \sum_{i=1}^{3} \sum_{j=1}^{5} c_{ij}\, x_{ij}. \tag{2.12}$$

Let us first find an initial trial solution which satisfies the constraints, and then inspect it for optimality. First find the lowest cost coefficient, in this case it is $c_{25} = 0$, then assign the maximum number of units, in this case 80. Now in the second row there are 20 more units available. Assign these to the lowest cost coefficients in that row, making sure to observe the column constraint. In this case we can assign all 20 to the fourth column. Then in this column look for the lowest cost coefficient and assign the remaining 60 of the fourth row. We proceed in this fashion through the matrix from row to column making assignments until all units are assigned. The assignment is as follows:

	r_j / a_i	W_1	W_2	W_3	W_4	W_5
		40	40	80	80	80
F_1	100	40	40	20	0	0
F_2	100	0	0	0	20	80
F_3	120	0	0	60	60	0

Now, let us test this solution for optimality by inspecting what would happen to this total if we reassigned one unit to a place where we now have a zero. For example, let us see what would happen if we made an entry at x_{35}. If we add one to x_{35}, we must subtract one from x_{25}, add one to x_{24} and finally subtract one from x_{34}. This shift would not violate the constraints and change the cost by

$$c_{35} - c_{25} + c_{24} - c_{34} = \Delta c \qquad (2.13)$$

$$4 - 0 + 1 - 1 = +4.$$

Thus, we see that this shift will increase the cost by 4 for every unit we ship from F_3 to W_5. By using this method, we can check the entire solution. If we find a shift which produces a negative total, then we would transfer as many units as possible through that shift.

Since c_{12} is the highest cost coefficient, let us look at shifting some units out of there. Let us try for a positive entry at x_{22}; then the shift would alternately add and subtract to x_{22}, x_{24}, x_{34}, x_{33}, x_{13} and x_{12} or the cost would change by

$$c_{22} - c_{24} + c_{34} - c_{33} + c_{13} - c_{12} \qquad (2.14)$$

$$2 - 1 + 1 - 2 + 2 - 5 = -3$$

This path is constrained by x_{24} (or x_{13}) which are 20; therefore we can transfer 20 units through this shift.

This new assignment is shown below:

	r_j / a_i	W_1	W_2	W_3	W_4	W_5
		40	40	80	80	80
F_1	100	40	20	40	0	0
F_2	100	0	20	0	0	80
F_3	120	0	0	40	80	0

This procedure could be carried out for every unoccupied place, but could become rather lengthy. The method of "shadow costs" allows us to quickly inspect a solution and determine if it is the minimum or not, and if not, what transfer to make. We will define shadow costs, u_i for every row and v_j for every column such that $u_i + v_j = c_{ij}$ for each non zero x_{ij}. Let us look at the first transfer we considered,

$$c_{35} - c_{25} + c_{24} - c_{34} = \Delta C \text{ or using} \tag{2.15}$$

the shadow costs for all but the zero x_{ij}, we have

$$c_{35} - u_2 - v_5 + u_2 + u_4 - u_3 - u_4 = \Delta C \tag{2.16}$$

$$c_{35} - (v_5 + u_3) = C \tag{2.17}$$

Thus, we see that if the shadow cost of the unoccupied cell does not exceed the true cost in that cell, then the total cost will increase by occupying that cell. Conversely, if the shadow exceeds the true cost, then an improvement can be realized.

$$\begin{aligned} c_{ij} &> u_i + v_j \quad \text{no shift} \\ c_{ij} &< u_i + v_j \quad \text{shift} \end{aligned} \tag{2.18}$$

In determining the values of the u_i's and v_j's, we assign an arbitrary value to one of them (usually set $u_1 = 0$) and work our way through the matrix

determining all other values of u_i and v_j, remembering only to evaluate u_i and v_j for the occupied cells. This can be seen more clearly by looking at the initial assignment.

40	40	20	0	0
0	0	0	20	80
0	0	60	60	0

We can now write the c_{ij}'s for the occupied cells, and then evaluate the u_i's and v_j's by starting with $u_1 = 0$.

v_j / u_i	2	5	2	1	0
0	2	5	2	x	x
0	x	x	x	1	0
0	x	x	2	1	x

With these values of u_i and v_j, we can determine the shadow costs for the unoccupied cells.

v_j / u_i	2	5	2	1	0
0	x	x	x	1	0
0	2	5	2	x	x
0	2	5	x	x	0

Here we see that $u_2 + v_2 > c_{22}$ and, therefore, we should improve the solution by transferring through that cell.

40	20	40	0	0
0	20	0	0	80
0	0	40	80	0

Again we must inspect this solution for optimality. Now writing the c_{ij}'s for the occupied cells, and evaluating the u_i's and v_j's, starting with $u_1 = 0$.

v_j / u_i	2	5	2	1	3
0	2	5	2	x	x
-3	x	2	x	x	0
0	x	x	2	1	x

And then the shadow costs for the unoccupied cells

v_j / u_i	2	5	2	1	3
0	x	x	x	1	3
-3	-1	x	-1	-2	x
0	2	5	x	x	3

Now we see that all the shadow costs are less than the true costs and therefore the solution must be a minimum.

The transportation problem, like the assignment problem can also be solved by the use of an algorithm, which is shown in Figure 2.21 and described below:

Algorithm For Transportation Problem

1. Subtract from each row its smallest element. Subtract from each column its smallest element. Pick a trial set of quotas by assigning them to zeros, subtracting appropriately from the discrepancies. Go to step 2.
2. Cover each column whose discrepancy is zero. If all columns are covered, the solution is complete; if not, go to step 3.

3. Look for an uncovered zero. If there is none, go to step 7. If one is found, go to step 4.

4. Prime this zero. Check the discrepancy of the row; if it is not zero, go to step 6. If it is zero, go to step 5.

5. Cover the row, and for each twice-covered essential zero, star the zero and uncover its column. Go to step 3.

6. There now exists a unique chain, as above. Find the smallest of the following numbers: the discrepancy of the row of the first primed zero in the chain; the discrepancy of the column of the last primed zero in the chain; the quota of each starred zero in the chain. This number is to be subtracted from each of these two discrepancies, and from the quota of every starred zero in the chain, and to be added to the quota of every primed zero in the chain. Now erase all primes and stars, uncover all rows, and go to step 2.

7. Find the smallest uncovered element in the matrix. Add this element to the covered rows and subtract it from the uncovered columns (or add it to the covered columns and subtract it from the uncovered rows) (or add it to the twice-covered elements and subtract it from the uncovered elements.) Do not change any stars, primes, or coverings. Go to step 3.

Note: "Discrepancies" are amounts to be shipped which have not yet been assigned. "Quotas" are amounts which have already been assigned to particular elements of the matrix (i.e., particular routes.) An essential zero is one whose quota is greater than zero.

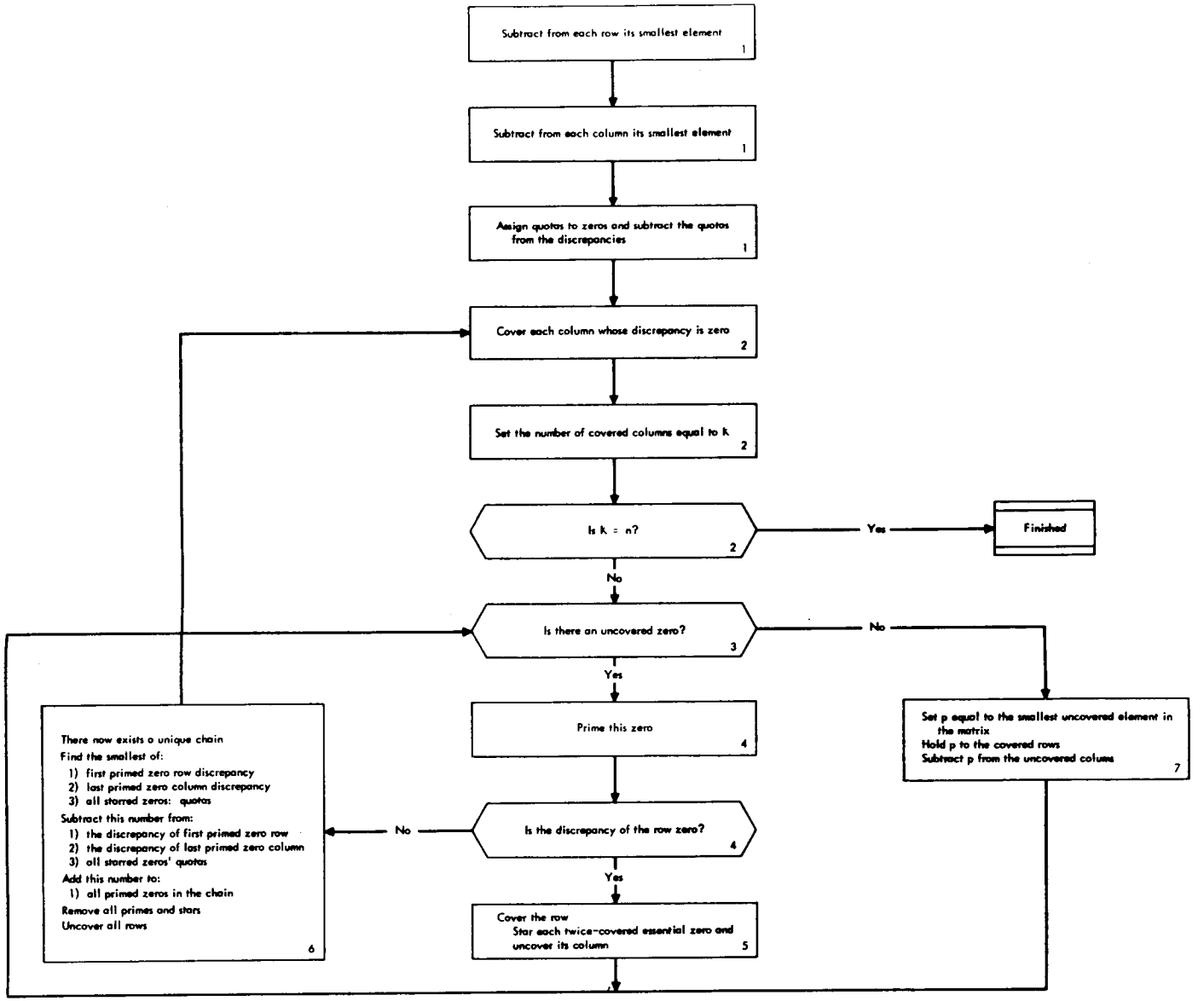


Figure 2.21. Algorithm for Transporting Problem

It is interesting to notice that a transportation problem is in reality a special form of the assignment problem. This implies that if in a transportation problem a_i and r_j are integers, all the x_{ij}'s must also be integers. To understand this observation, let us define both the assignment problem and then the transportation problem in the general case.

2.3 Assignment Problem Definition

An assignment problem is of the form that k men are to be assigned to k jobs and the cost of assigning each man to each job is uniquely defined, and we wish to find the minimum assignment of all men. This problem can be represented in a matrix form as shown in Figure 2.22.

	J_1	J_2	J_3	- -	J_j	- -	J_k
M_1	c_{11}	c_{12}	c_{13}	- -	c_{1j}	--	c_{1k}
M_2	c_{21}	c_{22}	c_{23}	- -	c_{2j}	- -	c_{2k}
M_3	c_{31}	c_{32}	c_{33}	- -	c_{3j}	- -	c_{3k}
M_i	c_{i1}	c_{i2}	c_{i3}	- -	c_{ij}	- -	c_{ik}
M_k	c_{k1}	c_{k2}	c_{k3}	- -	c_{kj}	- -	c_{kk}

Figure 2.22

c_{ij} is the cost of assigning M_i to J_i

y_{ij} is the assignment index, such that

$y_{ij} = 1$ if M_i is assigned to J_j, and

$y_{ij} = 0$ if M_i is not assigned to J_j

$$\sum_{j=1}^{k} y_{ij} = 1, \; i = 1, 2, 3, \text{---}k$$

$$\sum_{i=1}^{k} y_{ij} = 1, \; j = 1, 2, 3 \text{---}k$$

The cost function, C, to be minimized is

$$C_{min} = \sum_{i=1}^{k} \sum_{j=1}^{k} c_{ij} y_{ij}$$

2.4 Transportation Problem Definition

A transportation problem is of the form that there are m destinations and n sources, each source has a_i units available, and each destination has r_j units required. The cost of transporting from each source to each destination is uniquely defined, and we wish to find the minimum cost of transporting all units to the destinations. This problem can be represented in a matrix form as shown in Figure 2.23.

r_j / a_i		D_1	D_2	D_3	---	D_j	---	D_m
		r_1	r_2	r_3	---	r_j	---	r_m
S_1	a_1	c_{11}	c_{12}	c_{13}		c_{1j}		c_{1m}
S_2	a_2	c_{21}	c_{22}	c_{23}		c_{2j}		c_{2m}
S_3	a_3	c_{31}	c_{32}	c_{33}		c_{3j}		c_{3m}
S_i	a_i	c_{i1}	c_{i2}	c_{i3}		c_{ij}		c_{im}
S_n	a_n	c_{n1}	c_{n2}	c_{n3}		c_{nj}		c_{nm}

Figure 2.23

D_j is the jth destination

r_j is the number of units required at D_j

S_i is the i^{th} source

a_i is the number of units available at S_i

c_{ij} is the cost of shipping one unit from S_i to D_j

x_{ij} is the number of units shipped from S_i to D_j

If $\sum_{i=1}^{n} x_{ij} = r_j$, j = 1, 2, 3 --- m

and if $\sum_{j=1}^{m} x_{ij} = a_i$, $i = 1, 2, 3 \text{ --- } n$,

then $\sum_{i=1}^{n} a_i = \sum_{j=1}^{m} rj$

This last statement implies the number of units required is exactly equal to the number available. While this may not always be true in the original problem, any problem can be formulated in this manner by the additional artificial sources or destinations.

The cost function, C, to be minimized is

$$C_{min} = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} x_{ij}$$

Now, with these definitions, any transportation problem may be expanded into a large assignment problem as follows. Consider an m by n transportation problem where

$$\sum_{i=1}^{n} a_i = \sum_{j=1}^{m} r_j = k$$

This transportation problem can then be expanded into a k by k matrix which is composed of may subsets which are matrices with identical c_{ij}'s in each element and are a_i by r_j large ($i = 1, 2, \text{ --- } n$; $j = 1, 2 \text{ --- } m$). This problem is now similar to a standard assignment problem, and its optimum solution must contain values of $y_{ij} = 1$ or 0, where all the y_{ij}'s = 1 are independent. This large assignment problem can then be collapsed to the original transportation problem, and the x_{ij}'s of the transportation will be the sum of the y_{ij}'s of the assignment problem for each constant c_{ij} subset of the matrix. Since the y_{ij}'s of the assignment problem can only have values 1 or 0, integer numbers, then their sums, the x_{ij}'s of the transportation problem, must also be integers.

3 DECISION THEORY

The subject of decision theory may have more faces and interpretations than one can imagine. We will concern ourselves with some of the more practical applications of decision theory to gain some insight. Let us consider the following situation:

You are sitting on a park bench on a nice cool Sunday afternoon watching whatever seems to be the thing to watch. Then along comes a rather average young man; you hardly take notice of him, until he sits down next to you. Then you notice he has an ice cream cone in his hand. He offers you the ice cream cone. You have just finished eating and without too much thought reply, "no, thank you". You just don't feel like eating an ice cream cone. Now, in refusing the cone, did you make a decision? No, not really, you simply reacted.

Now the stranger reaches into his pocket with his other hand and pulls out a small hand gun. He looks at you, places the gun in your side and says, "I think you want my ice cream cone". What do you do? You accept the ice cream cone rather quickly and thank him. Now, have you made a decision? No, you had no other alternative but to accept. At least if you are a rational person, you had no other alternative. You could have chosen to get shot, but that is hardly rational.

So there you are about to eat your ice cream cone, still not having made any decisions, when another man sits down along side of you. This second man whispers in your ear not to eat the ice cream cone. He claims that he knows the first man with the gun in your side and that this man goes around the park on Sunday afternoon passing out poisoned ice cream cones. He claims that this man is a little crazy.

Now you have a problem and, now, you have to make a decision.

You also start asking yourself some questions which you must answer before you can make any decision. Is the ice cream cone really poisoned? What kind of poison? Who is really crazy, the first man, the second man, both men or neither? Does the first man have a real gun? Is it really loaded? Will he really shoot?

Now, what are your choices? You can eat the ice cream cone, hope nothing happens, get up, say goodbye and go to the nearest hospital to have your stomach pumped; you can eat it and take your chances; you can refuse to eat it and take your chances; or you can drop the ice cream cone and run for your life and hope the man is a poor shot.

It is this type of decision making which is to be discussed. Let us first define some of the major elements or key points of a decision.

1) Problem

There must be a problem before we can talk about making a decision. There must be several choices at hand and the correct choice is not obvious. We must consider a decision as an irrevocable commitment of resources.

2) Uncertainty

To be concerned with a decision, there must be uncertainty as to what the actual outcome will be. If the outcome is determined, then no real decision is required.

3) Probability Theory

The fact that there is uncertainty involved implies that probability theory must play a strong part in decision theory.

4) State of Mind.

While the theory of probability is underlying to the entire approach of making a decision, we must use the probability as our state of mind about the situation. The probability of an event occurring is a measure of our belief that it will occur.

5) Experience

The probability we associate with an event occurring is dependent upon our experience.

6) Value

We must assign a value to each possible outcome of a situation. The value of an outcome must not be confused or influenced by the probability of the outcome occurring.

7) Risk Criteria

We must determine how much risk we are willing to take before we make the final decision. Do we want to maximize the expected value and reduce the probability of obtaining it, or do we want to minimize the the probability of getting nothing?

8) Future

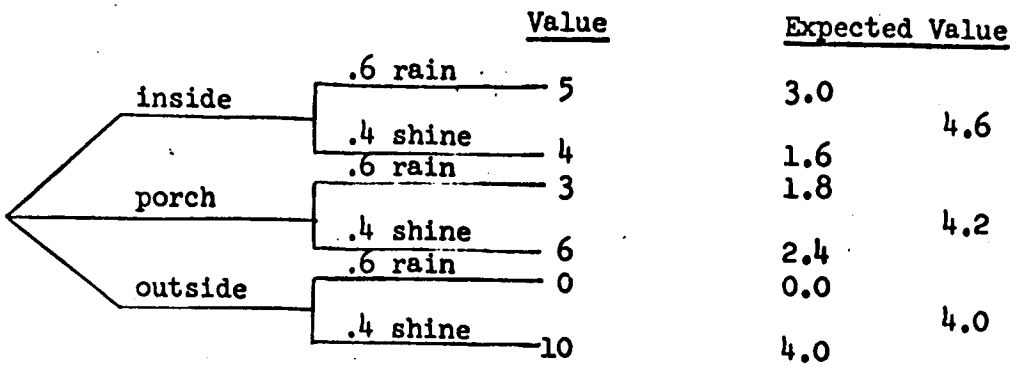
A decision must be influenced by the future. When a decision is made, its outcome is dependent on things that happen subsequently.

9) Outcome

It is very important to realize that good decisions can have good outcomes or bad outcomes, and that bad decisions can have good outcomes or bad outcomes.

It is interesting to note that Decision Theory is highly dependent on Probability Theory, yet Probability Theory is 350 years old; whereas, Decision Theory is only 20 years old.

Let us look at a problem in which a girl cannot decide where to have her wedding reception: inside, outside, or on the porch. The real problem is rain or shine. Let's look at all possible outcomes and assign some value numbers to them.



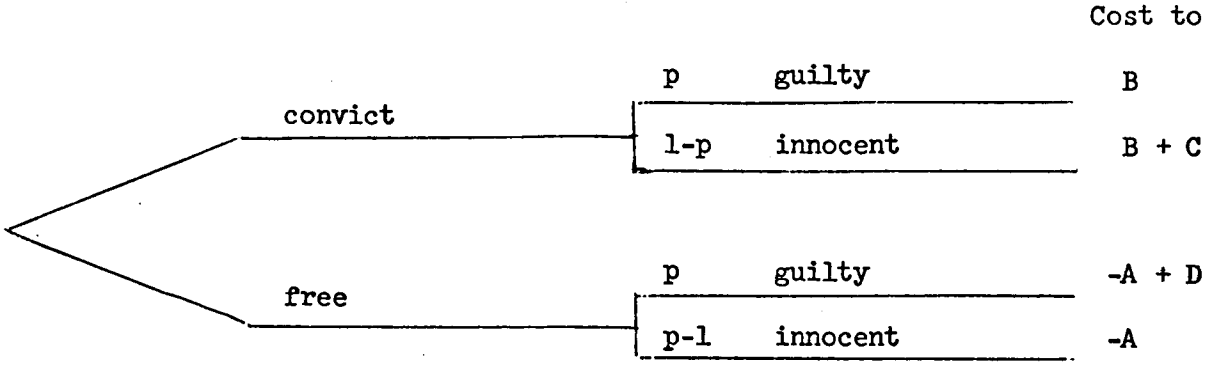
We can now determine the expected value of each decision (i.e. inside, outside, porch) which is the sum of the values times their probabilities.

	Expected Value	
Probability of rain	.4 to .6	.5 to .5
Inside	4.6	4.5
Porch	4.2	4.5
Outside	4.0	5.0

The above table shows how a small change in the probability of rain can change the expected value.

In addition to considering the Expected Value we must also consider our willingness to take a risk. With a 50 - 50 chance of rain our expected value is the highest if the party is outside, but 50% of the time we have a flop (i.e. value 0). Whereas, if the party is inside the expected value is 4.5 and the true value is never below 4.

Let us consider another problem in which a judge must decide whether to free or convict a man, given some probability, p, that the man is guilty and 1-p that he is innocent.



A	Gain to society of freeing the man	$ 7,000
B	Cost of keeping a man in jail	$ 2,000
C	Cost of convicting an innocent man	$100,000
D	Loss to society of freeing a guilty man	$10,000

Now, the Expected Values of each choice are:

Convict $pB + (1-p)(B+C)$

Free $p(D-A) - (1-p)A$

Therefore the judge should convict if $pB + (1-p)(B+C) > p(D-A) - (1-p)A$

$pB + B + C - pB - pC > pD - pA - A - pA$

$pC + pD > B + C + A$

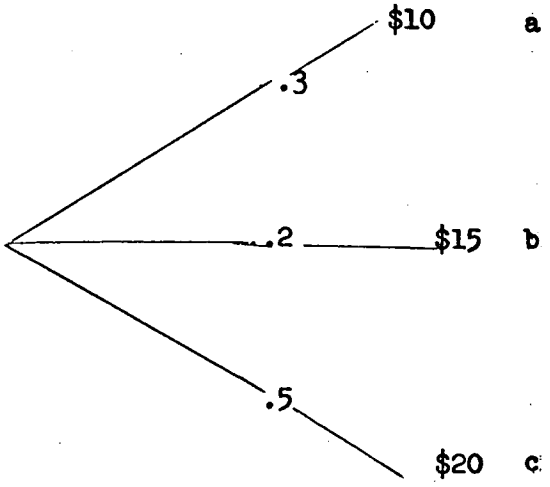
$$p > \frac{A+B+C}{C+D}$$

In the case shown, p must be

$$p > \frac{7,000 + 2,000 + 100,000}{100,000 + 10,000}$$

$$p > \frac{109}{110} = .991$$

We will use the concept of lotteries to illustrate some of the essential components of making a decision. A lottery may be considered as a game-of-chance in which one must pay a given amount to play the game with a fixed (sometimes known, sometimes unknown) probability of winning a defined prize. We will use lotteries which have several well-defined outcomes, each of which has a fixed probability or occurring. These lotteries can be represented as shown below:



This lottery has three possible outcomes (a, b or c); these outcomes have prizes of $10, $15 and $20 respectively, and probabilities of .3, .2 and .5 respectively. Note that the sum of the probabilities is equal to one.

We will now discuss five properties of lotteries: orderability, continuity, substitutiability, monotonicity, and decomposibility.

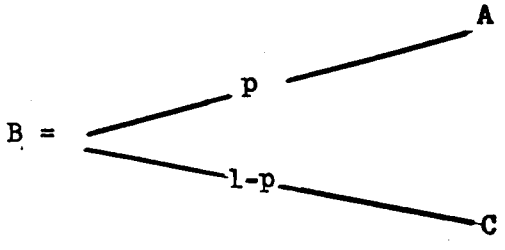
Orderability says that either lottery A is better than lottery B or lottery B is better than lottery A.

$$\text{Orderability} = A < B \text{ or } a > B$$

Continuity says that if lottery A is better than lobbery B which is better than lottery C.; then there is some probability, p, such that

$$B \backsim [p, A; (1-p)\ C]$$

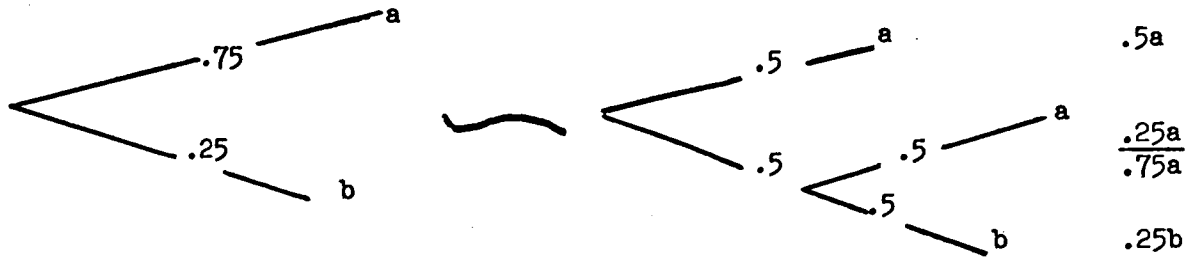
$B \backsim \tilde{B}$, is the certainty equivalent



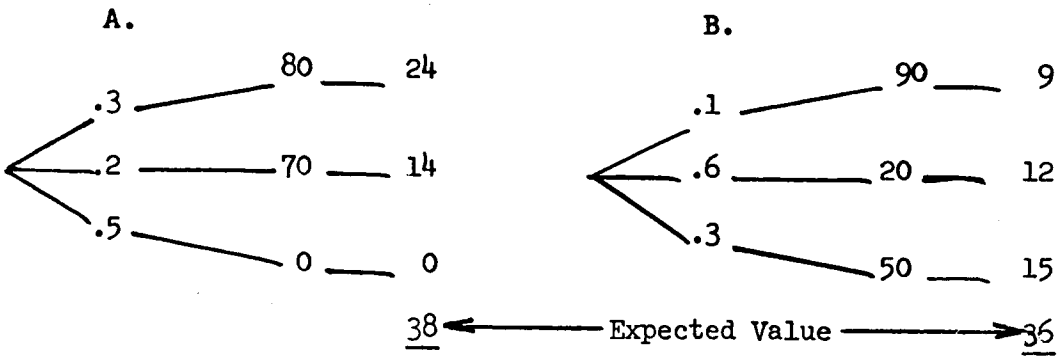
Substitutability says that if two lotteries are equal, then either can be chosen.

Monotonicity says that if lottery A is better than lottery B, then lottery B should be chosen.

Decomposability says that the following two lotteries are equivalent:

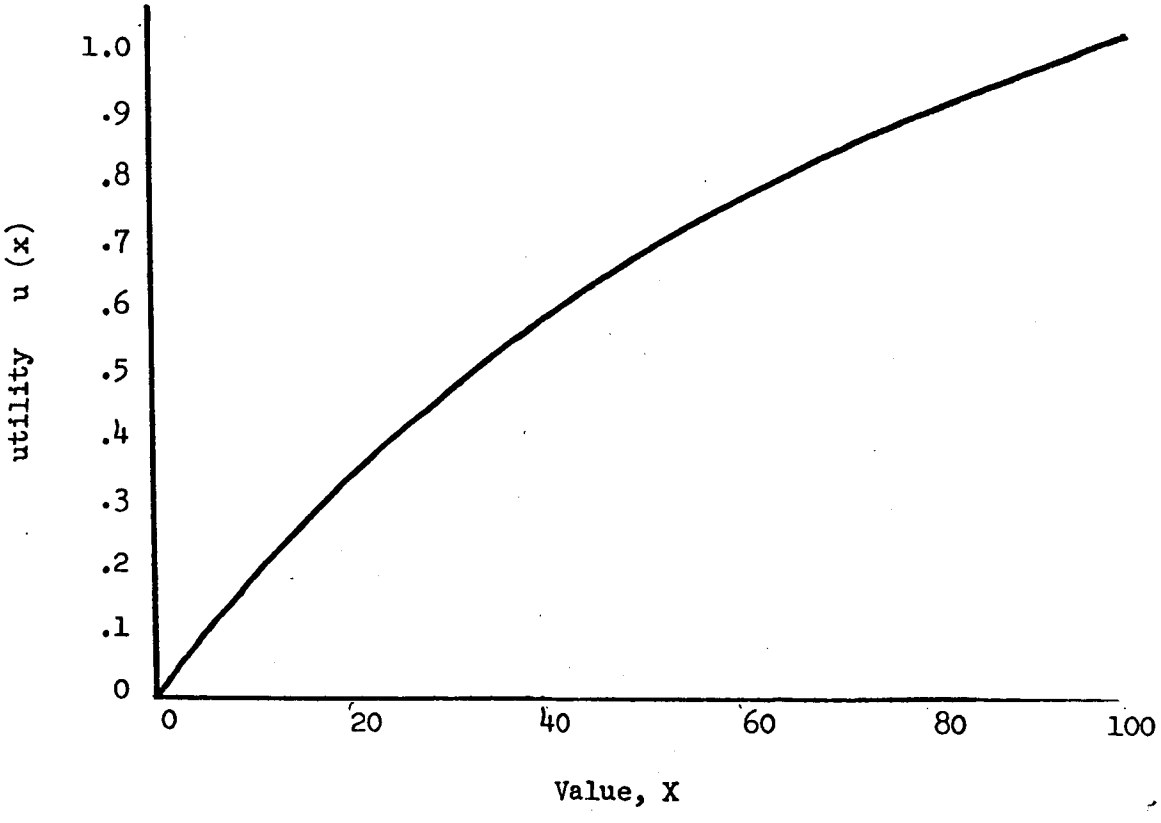


Let us look at two lotteries:

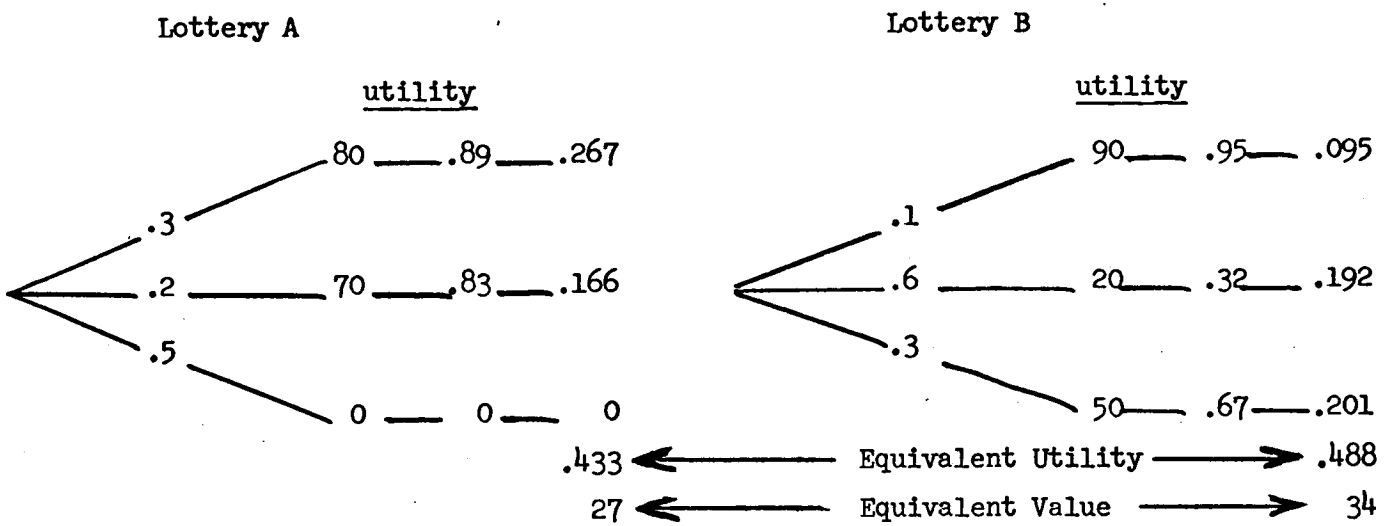


From the Expected Value, only lottery A is the proper choice. It should be noted that in lottery A, 50% of the time we get nothing, while in lottery B, 40% of the time we get 50 or better.

Now, to consider the utility of these two lotteries, it is necessary to determine the risk character. Let us assume we are a risk averter with a utility function: $u(x) = 4/3 \left[1 - (\frac{1}{2})^{\frac{x}{50}}\right]$



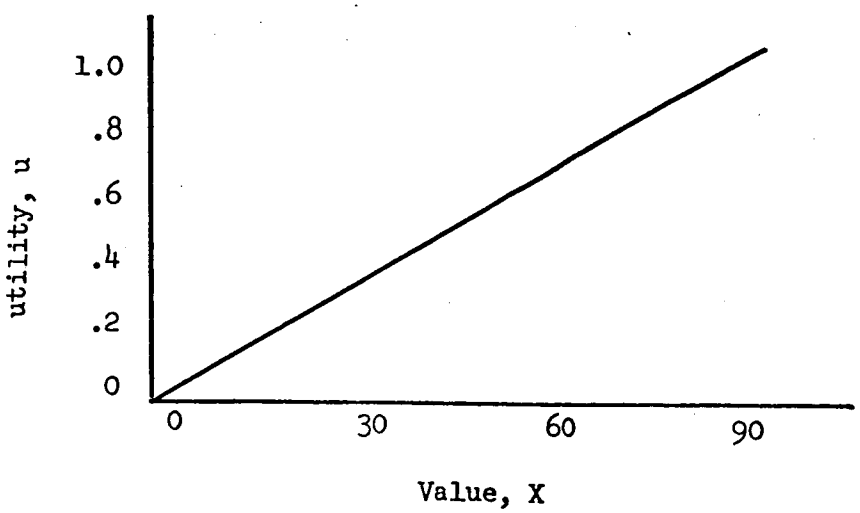
We can now determine the utility of both lotteries and then their value (for this risk averter).


Lottery A
utility
80 .89 .267
.3
.2
70 .83 .166
.5
0 0 0
.433
27
Lottery B
utility
90 .95 .095
.1
.6
20 .32 .192
.3
50 .67 .201
Equivalent Utility
.488
Equivalent Value
34

It is now evident that the Equivalent Value is higher for Lottery B than for lottery A, but both are lower than the Expected Value. In other words, for lottery A, with an expected value of 38, we would pay 27, and for lottery B, with an expected value of 36, we would pay 34.

Risk Indifferent

A risk indifferent person has a utility function with a one to one relationship between utility and value



slope = 1

With the advent of the space age the field of astrodynamics has grown rapidly. Most of the tools which are used today were essentially developed by Newton and others, out of the basic discoveries of men like Kepler, Brahe and Copernicus. This work of Newton, later embelished by such mathematicians as Lagrange and Euler, composes what is classically called Celestial Mechanics. Astrodymanics, as we shall briefly study it, is engineering or practical application of Celestial Mechanics to the contemporary problems of space vehicles, exclusive of conventional aerodynamics and booster propulsion theory.

In dealing with the trajectories of an artificial sattelite or an interplanetary spacecraft, it is convenient to consider the general problem as a set of problems each of which can be considered as a two body problem. In the case of a trajectory from the Earth to a target planet (e.g. Mars), this is done by first considering the Earth as the center of the coordinate system, then the Sun, then Mars. By using this approach, we can simplify what could be an extremely difficult problem. This simplification lets us get a close answer to the real problem and a quick understanding of the situation.

Let us look first at the two body problem in its general form, then at the coordinate systems of interest and finally at some specific problems.

In astrodynamics, when we talk about the two-body problem we are restricting our thinking to motion of one (relatively small) body about another (relatively large) body. The second body is usually considered as the central force field, and is used as the center of the coordinate system in which the motion of the smaller body is described. For example, one can consider the motion of the Earth about the Sun. In this "two-body problem", the Sun is the center force field, and the origin of an orthogonal coordinate system in which the Earth's motion can be described. In this system, a surprisingly accurate

description of the Earth's motion can be developed without considering the effects, or purturbations of the other planets on the Earth.

There exists a fundamentally important relationship which uniquely describes the motion of a body in orbit about another body, called the vis-viva integral. This integral is commonly seen in the following form:

$$C_3 = V^2 - \frac{2GM}{R}$$

where: $GM = 3.9 \times 10^5 \ \frac{km^3}{sec^2}$ for Earth

R = the radius to the vehicle from the center of the Earth

V = velocity of the vehicle at a distance R

C_3 = twice the total geocentric energy per unit mass in km^2/sec^2. C_3 is also the square of the hyperbolic excess velocity.

Another form of the vis-viva integral which is extremely useful and simple is the following dimensionless form.

$$\dot{s}^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right)$$

All quantities are used in a dimensionless form, more will be said about this later.

r is the distance of the vehicle from the center of the system

$\dot{s}$ is the speed of the vehicle at a distance r

a is the semi-major axis of the conic of the vehicle.

It can be seen by inspection of this equation that it is the sum of the total potential and kinetic energy of the system. Specifically;

$\dot{s}^2$ corresponds to the kenetic energy

$\frac{2\mu}{r}$ corresponds to the potential energy

$-\frac{\mu}{a}$ corresponds to the total energy and is constant.

The study and fundamental understanding of this vis-viva integral is a most powerful tool. Its use in the conceptual design or system engineering of space ventures is almost unlimited. We will concern ourselves mainly with the use of this equation and forego its formal development.

Kepler's first law states, "The orbit of each planet is an ellipse with the sun at a focus." Newton later expanded this law to state that in all two body problems the motion under a central force field results in conic sections. The conic sections and some of their important constants and forms of the vis-viva integral are:

Circle

$e = 0$

$a = r$

$$\dot{s}^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right)$$

$$\dot{s}^2 = \frac{1}{r}$$

Ellipse

$0 > e > 1$

$a > 0$

$$\dot{s}^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right)$$

Parabola

$$e = 1$$

$$a = \infty$$

$$\dot{s}^2 = \mu \left(\frac{2}{r} - \frac{1}{a}\right)$$

$$\dot{s}^2 = \mu \left(\frac{2}{r}\right)$$

Hyperbola

$$e > 1$$

$$a < 0$$

$$\dot{s}^2 = \mu \left(\frac{2}{r} - \frac{1}{-a}\right)$$

$$\dot{s}^2 = \mu \left(\frac{2}{r} + \frac{1}{a}\right)$$

In astrodynamics and astronomy it is often useful, and more accurate, to use a dimensionless system; that is, the numbers which are actually used in the computation have no dimensions and are actually ratios to well known parameters. If we consider the quantities of length, speed, and mass we have the following basics to use:

Length

Geocentric

In considering systems in which the Earth is the central force field all linear dimensions are expressed in terms of the Earth's radius. Then in this system the distance to the surface of the Earth is:

$$r_{\oplus} = 1 = \frac{3957 \text{ mi}}{3957 \text{ mi}}$$

where we will assume the radius of the Earth is 3957 miles. The distance to the moon is, $r_{\leftmoon} = 60.3706$.

Heliocentric

In considering systems in which the Sun is the central force field, all linear dimensions are expressed in terms of the semi-major axis of the Earth's orbit about the Sun. This distance is one astronomical unit (1 a.u.) or approximately 92.90×10^6 mi. In this system, the distance to all the planets is:

Mercury	0.3871
Venus	0.7233
Earth	1.0000
Mars	1.5237
Jupiter	5.2028
Saturn	9.5388
Uranus	19.1820
Neptune	30.0577
Pluto	39.5177

Speed

The speed is in terms of the satellite speed at a unit distance.

Geocentric

In this system the speed is the satellite speed at one Earth radius. The actual speed is 7.905 km/sec, 26,000 ft/sec, 4.912 mi/sec ($\simeq$ 5 mi/sec).

Heliocentric

In this system, the speed is the satellite speed at 1 a.u., or the speed of the Earth in its own orbit. The actual speed is 29.6 km/sec, 96,700 ft/sec, 18.6 mi/sec.

Mass

The mass is expressed in terms of the most massive body in the system (i.e. the central body). In the vis-viva integral

$$\dot{s}^2 = \mu \left(\frac{2}{r} - \frac{1}{a}\right)$$

is the sum of the two masses in the system in dimensionless form, $m_1 = 1$ and usually $m_1 > m_2$, therefore

$$\mu = m_1 + m_2 \simeq 1$$

Some useful Earth mass ratios are

Sun	331,950
Moon	0.012
Mercury	0.05
Venus	0.81
Earth	1.00
Mars	0.11
Jupiter	318.4
Saturn	95.3
Uranus	14.5
Neptune	17.2

Let us use these concepts and determine the altitude and speed of a 24 hour synchronous satellite in a circular orbit.

First we can determine the semi-major axis by the use of Kepler's third law.

$$\left(\frac{P_\Delta}{P_\oplus}\right)^2 = \left(\frac{a_\Delta}{a_\oplus}\right)^3$$

Where the sub Δ refers to the satellite and the sub $\oplus$ refers to the Earth. Now since $a_\oplus = 1$, we can rewrite this expression

$$a_\Delta = \frac{P_\Delta}{P_\oplus}^{2/3}$$

Now $P_{\oplus}$ is the period of a satellite at one Earth radius or

$$P_{\oplus} = \frac{2 \cdot \pi \cdot 3957}{5 \cdot 60 \cdot 60}$$

$P_{\oplus}$ = 1.38 hr. and

P_{Δ} = 24 hr then

$$a_{\Delta} = \left(\frac{24}{1.38}\right)^{2/3}$$

a_{Δ} = 6.8 now the altitude, h, in miles is

$h = (a_{\Delta} - 1)\ 3957$

h = 23,000 miles

Now the speed of the satellite is

$$\dot{s}^2 = \mu \left(\frac{2}{r} - \frac{1}{a}\right)$$

but for a circular orbit r = a, therefore

$$\dot{s}^2 = \left(\frac{1}{r}\right)$$

where $r = a_{\Delta} = 6.8$

$\mu = 1$

$$\dot{s}^2 = 1 \left(\frac{1}{6.8}\right)$$

$\dot{s}^2 = .147$

$\dot{s}^2 = .384$

Then the speed in miles/sec, u, is

$u = \dot{s}^2 \cdot 5$

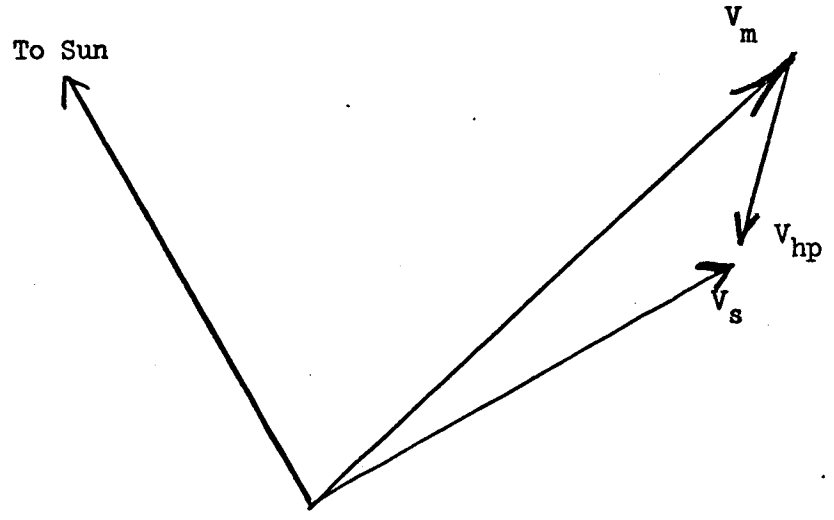
$u = .384 \cdot 5$

u = 1.92 miles/sec

4.1 Planetary Approach

The approach phase of planetary missions interacts with all other phases including launch, transit, communication distance to Earth and the flight time. While we will be discussing the problems associated with Mars, the general concepts are applicable to other planets with only minor modifications. To make the problems more tangible, we will consider the specifics of the Mars approach geometry as it will be during a 1971 opportunity.

The approach geometry at Mars is mainly determined by the magnitude and direction of the aerocentric hyperbolic excess velocity. This velocity is the vectoral difference between the heliocentric velocity of Mars and the heliocentric velocity of the spacecraft at the Mars encounter, neglecting the gravitational influence of Mars on the spacecraft. This relationship is shown in Figure 4.1.



V_m - heliocentric velocity of Mars

V_s - heliocentric velocity of the Spacecraft

V_{hp} - hyperbolic excess velocity

Figure 4.1

If we assume a simple coplaner Hohmann transfer between the Earth and Mars, we can obtain a quick estimate of the minimum value of V_{hp}. Using the geometry shown in Figure 4.2.

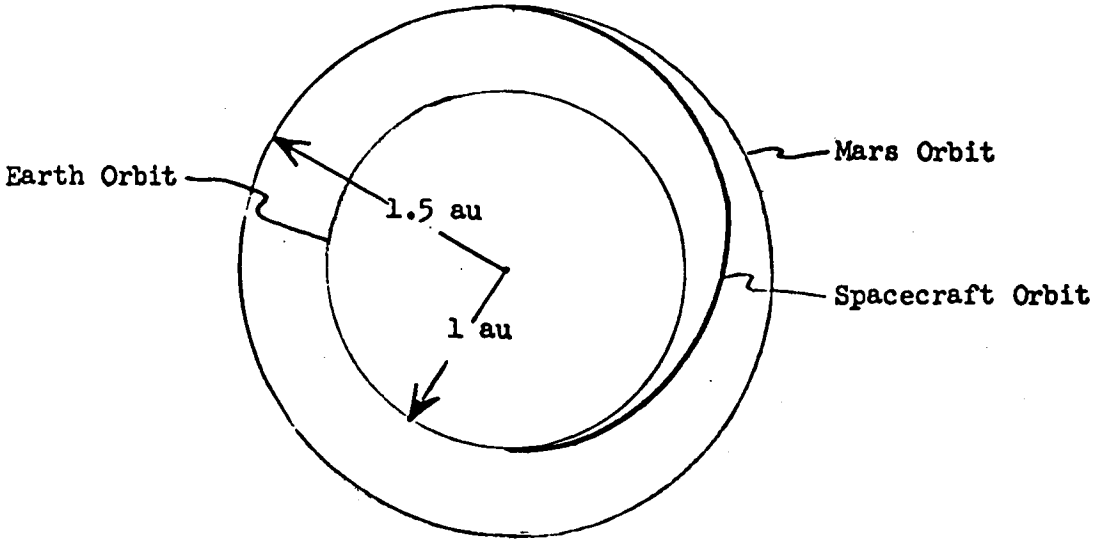


Figure 4.2

First computing the velocity of the spacecraft v_s we have

$$\dot{s}^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right)$$

$$a = \tfrac{1}{2}(1 + 1.5)$$

$$a = 1.25$$

$$r = 1.5$$

$$\mu = 1$$

$$\dot{s}^2 = 1\left(\frac{2}{1.5} - \frac{1}{1.25}\right)$$

$$\dot{s} = .73$$

Now the heliocentric velocity of the Earth is approximately 30 km/sec, therefore

$$v_s = 30 \times .73$$

$$v_s = 21.90 \text{ km/sec}$$

and the heliocentric velocity of Mars is

$$\dot{s}^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right)$$

$r = a = 1.5$ then

$$\dot{s}^2 = 1\left(\frac{1}{1.5}\right)$$

$s = .815$ and then

$v_m = 30 \times .815$

$v_m = 24.45$ km/sec

Now, since we assumed a Hohmann transfer

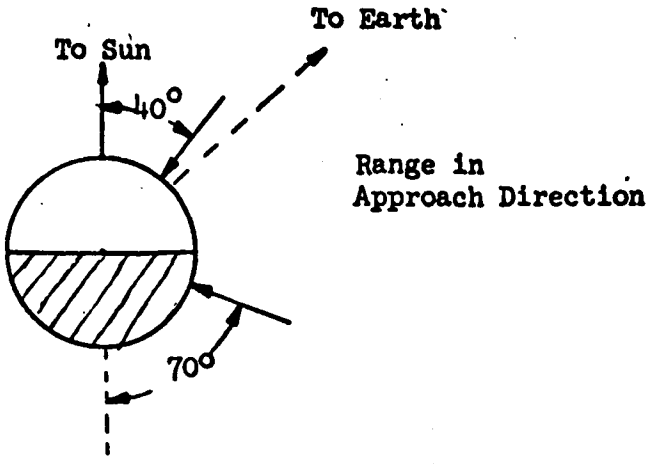
$v_{hp} = v_m - v_s$

$v_{hp} = 24.45 - 21.90$

$v_{hp} = 2.55$ km/sec

It must be noted that this value of the hyperbolic excess velocity is a minimum. Assuming the orbits are circular, the transfer is co-planer and Hohmann. The actual minimum is 2.82 km/sec for 1971 and 2.40 km/sec for 1973. By inspection of the vis-viva integral, it can be seen that the spacecraft velocity at Mars will always be less than the planet's velocity, since the value of r is identical and the semi-major axis of the transfer will always be less than the semi-major axis of Mars, for transfers that are reasonably close to optimum.

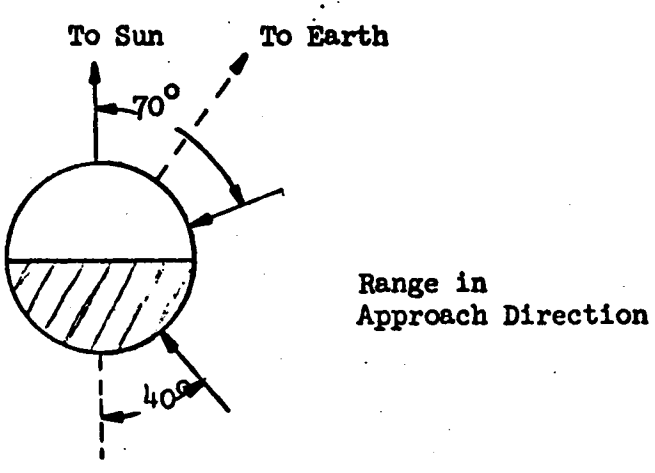
Two basic types of transfer trajectories must be considered, Type I and Type II. Type I trajectories have heliocentric transfer angles less than 180°, where the heliocentric transfer angle is measured from the position of the Earth at launch to the position of Mars at encounter. The Type I trajectories generally approach Mars from the lighted side, see Figure 4.3.



Type I Approach Direction

Figure 4.3

Type II trajectories have heliocentric transfer angles greater than 180°, and approach Mars generally from the dark side, see Figure 4.4.



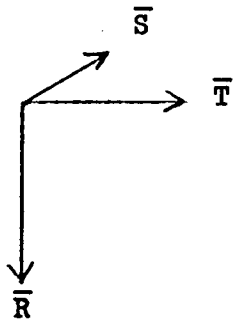
Type II Approach Direction

Figure 4.4

Some of the science experiments, television for example, are profoundly influenced by whether the approach is from the lighted side of the planet or from the dark side.

To understand and visualize the entire geometry problem about Mars better, we will define a Mars centered coordinate system. This system will be a

right handed, three dimension, orthogonal system, see Figure 4.5.



Approach Coordinate System

Figure 4.5

This system is composed of three unit vectors, $\bar{R}$, $\bar{S}$, $\bar{T}$ such that $\bar{R} = \bar{S} \times \bar{T}$. $\bar{T}$ is parallel to the ecliptic, $\bar{S}$ is parallel to the direction of the hyperbolic approach asymptote and $\bar{R}$ completes the system. It is important to understand this system in order to comprehend the entire approach problem.

Within this coordinate system, we need to know the location of the Earth and the Sun because we must communicate with the Earth and derive solar power from the Sun. We will first define an angle, ZAP, the angle between the Mars - Sun vector at encounter and the hyperbolic excess velocity vector. This angle is close to the Mars - spacecraft - Sun angle a few days before encounter. It should be noticed that if ZAP is less than 90°, the approach is from the dark side; for ZAP greater than 90°, the approach is from the lighted side. Another important angle to define is ETS, which is the angle measured clockwise from the $\bar{T}$ axis to the negative projection of the Mars - Sun vector onto the $\bar{R}$ - $\bar{T}$ plane. These two angles are shown in Figure 4.6.

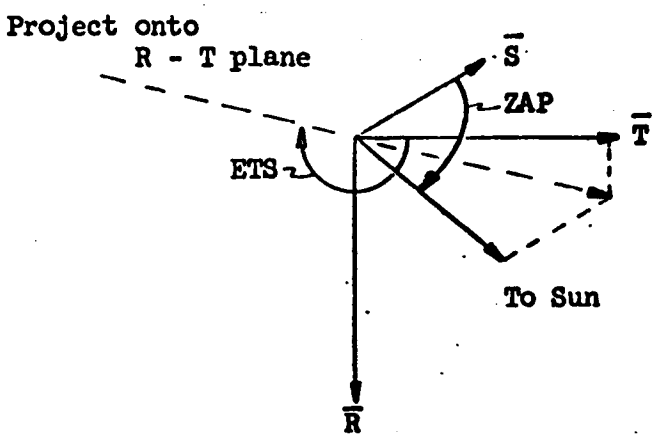


Figure 4.6

A similar pair of angles ZAE and ETE are defined for the Earth. ZAE is the angle between the Mars - Earth vector and the hyperbolic excess velocity vector. ETE is the angle measured clockwise from the T axis to the negative projection of the Mars Earth vector onto the R - T plane. The reason for measuring these angles (ETE and ETS) to the negative projections will become apparent.

To determine where the spacecraft actually flys by the planet in this coordinate system, we will define a point (the aiming point) in the R - T plane (the aiming plane) where the hyperbolic approach asymptote passes through that plane. The aiming point can then be defined by the vector $\bar{B}$ which has magnitude $|\bar{B}|$ and orientation θ to the T axis or by the components of $\bar{B}$, ($\bar{B} \cdot \bar{T}$ and $\bar{B} \cdot \bar{R}$) as shown in Figures 4.7 and 4.8.

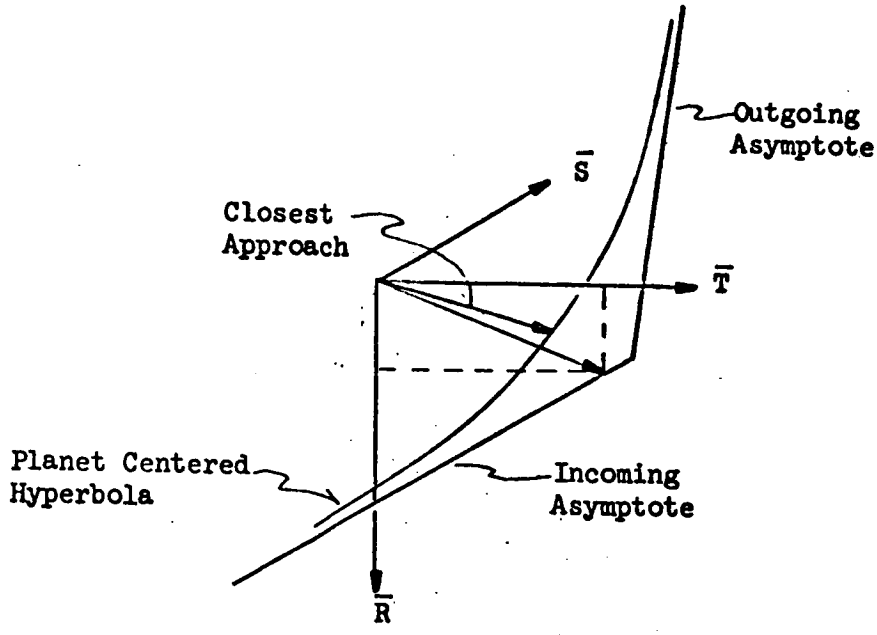
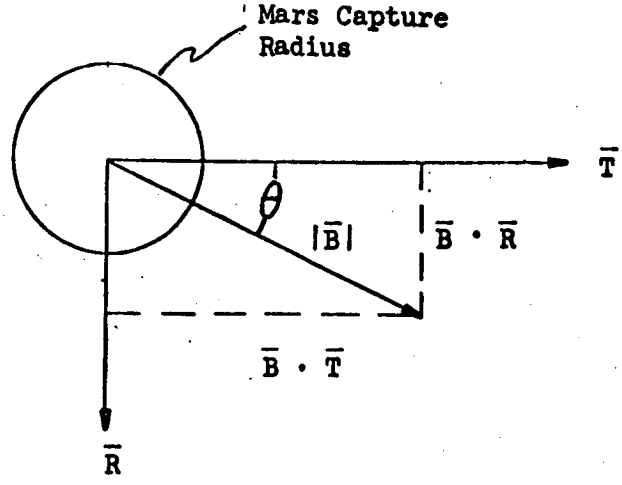


Figure 4.7

For clarity the aiming plane is shown in Figure 4.8,



Aiming Plane

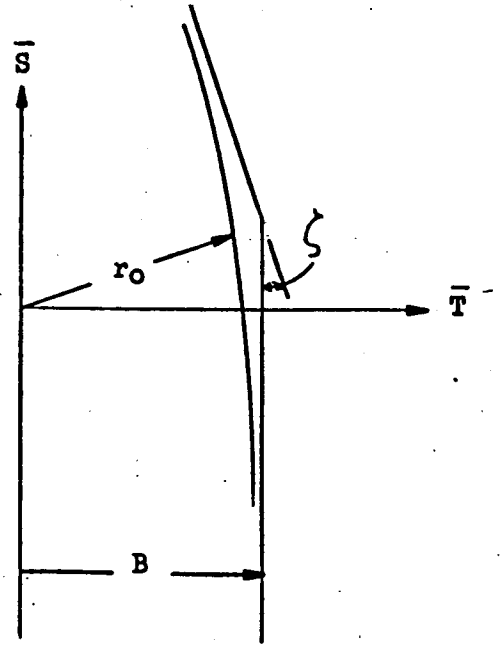
Figure 4.8

Now, with this definition, the reason for measuring ETE and ETS to the negative projections becomes apparent: if θ = ETE, then the Earth as seen by the spacecraft will be occulted and if θ = ETS, the Sun as seen by the spacecraft

will be occulted. The time after encounter at which these occultations will occur depends on the magnitude of B and the hyperbolic excess velocity. In the case of an orbiter, Earth occultation will occur at ETE and ETE + π , similiarly for ETS.

Now with this concept of targeting or aiming the approach asymptote in the R - T plane at a massless planet, it becomes easy to transfer from a heliocentric orbit to an areocentric orbit. This concept also allows the approach phase to be treated independently from the interplanetary phase. We can now investigate what the near planet geometry is in reality given a miss parameter and the hyperbolic excess velocity, v_∞ .

The first parameter of interest is the radius of closest approach, r_o. This is the shortest perpendicular distance from the actual flyby trajectory to the center of the planet (Figure 4.9). It should be noted in Figure 4.9 that $\overline{T}$



Miss Parameter B, Closest Approach r_o

Figure 4.9

is in the plane of the paper for $\theta = 0$ only. The relationship between r_o and B is

$$B = r_o \sqrt{1 + \frac{2\,GM}{r_o v_\infty^2}}$$

where for Mars

$$GM = 4.298 \times 10^4 \text{ km}^3/\text{sec}^2$$

this is shown in Figure 4.10 for different values of v_∞. Figure 4.10 actually has the altitude of periapsis versus the miss parameter, where $r_o = R_p + h_p$, where h_p is the altitude of periapsis.

The bending angle ζ as shown in Figure 4.9 is plotted in Figure 4.11, again versus miss parameter, for different values of v_∞.

The entry angle, θ_e, and range angle, ε, are two important parameters in considering the atmospheric entry problems. These parameters are shown in Figure 4.12 and plotted in Figure 4.13.

In addition to changing the shape of the approach trajectory, the gravitational effect of the planet on the spacecraft also increases the hyperbolic excess velocity, v_h. The hyperbolic excess velocity is v_∞ at very large distances from the planet ($R = \infty$) and v_e at the upper atmosphere. This entry velocity is shown in Figure 4.14. The hyperbolic excess velocity is related to v_∞ as shown below

$$v_h^2 = v_\infty^2 + \frac{2\,GM}{R}$$

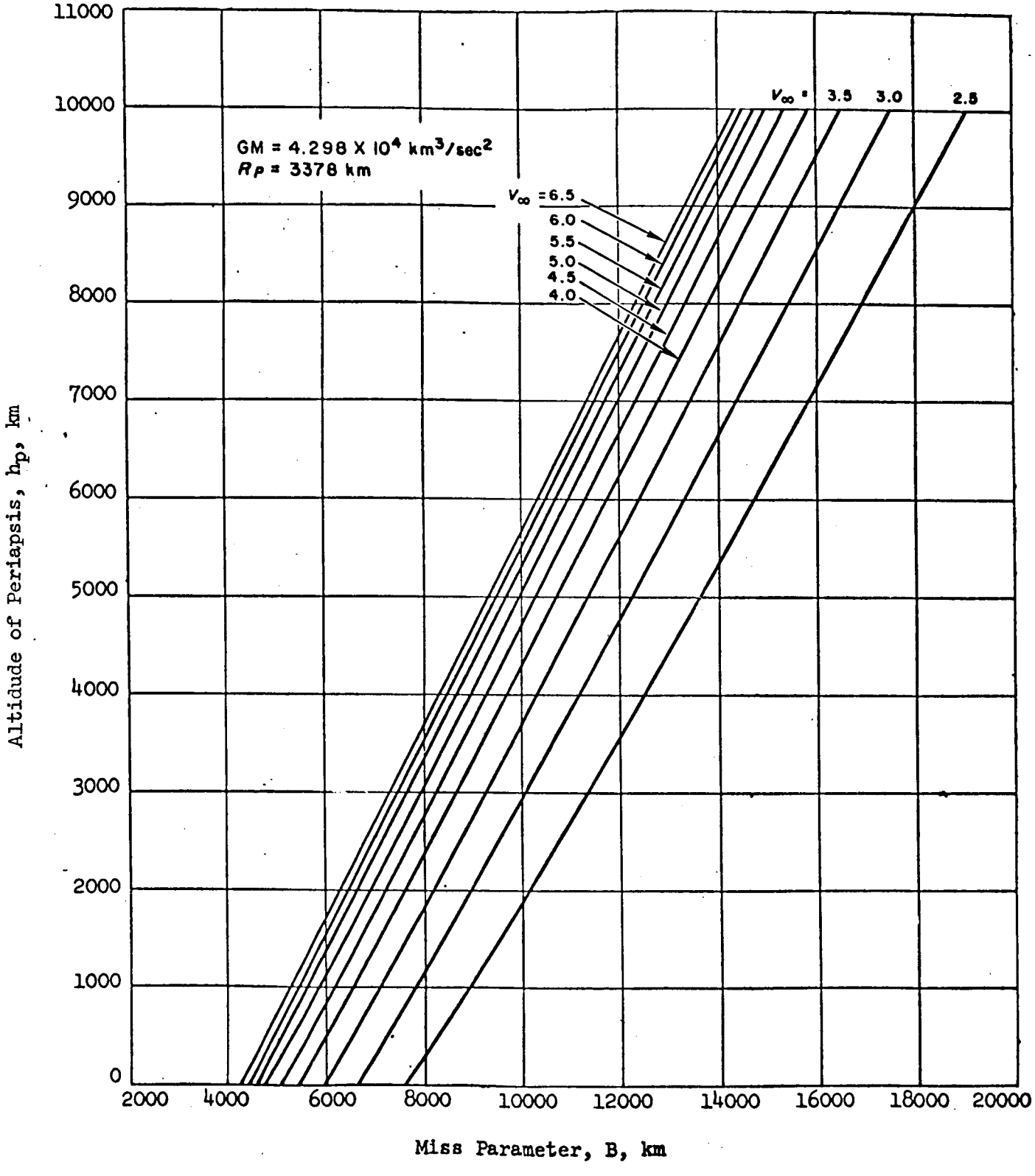


Figure 4.10. Altitude of Periapsis Versus Miss Parameter

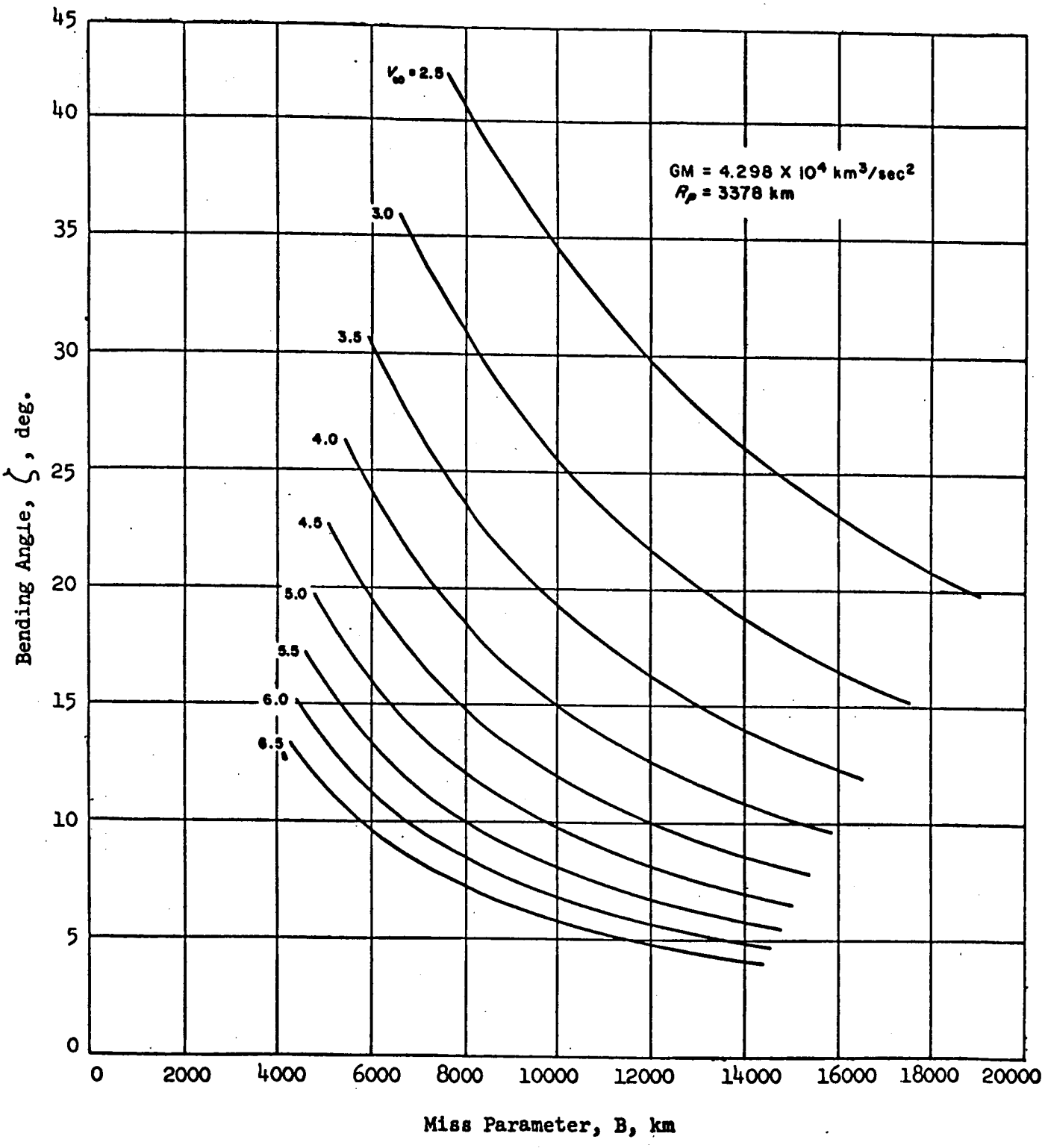
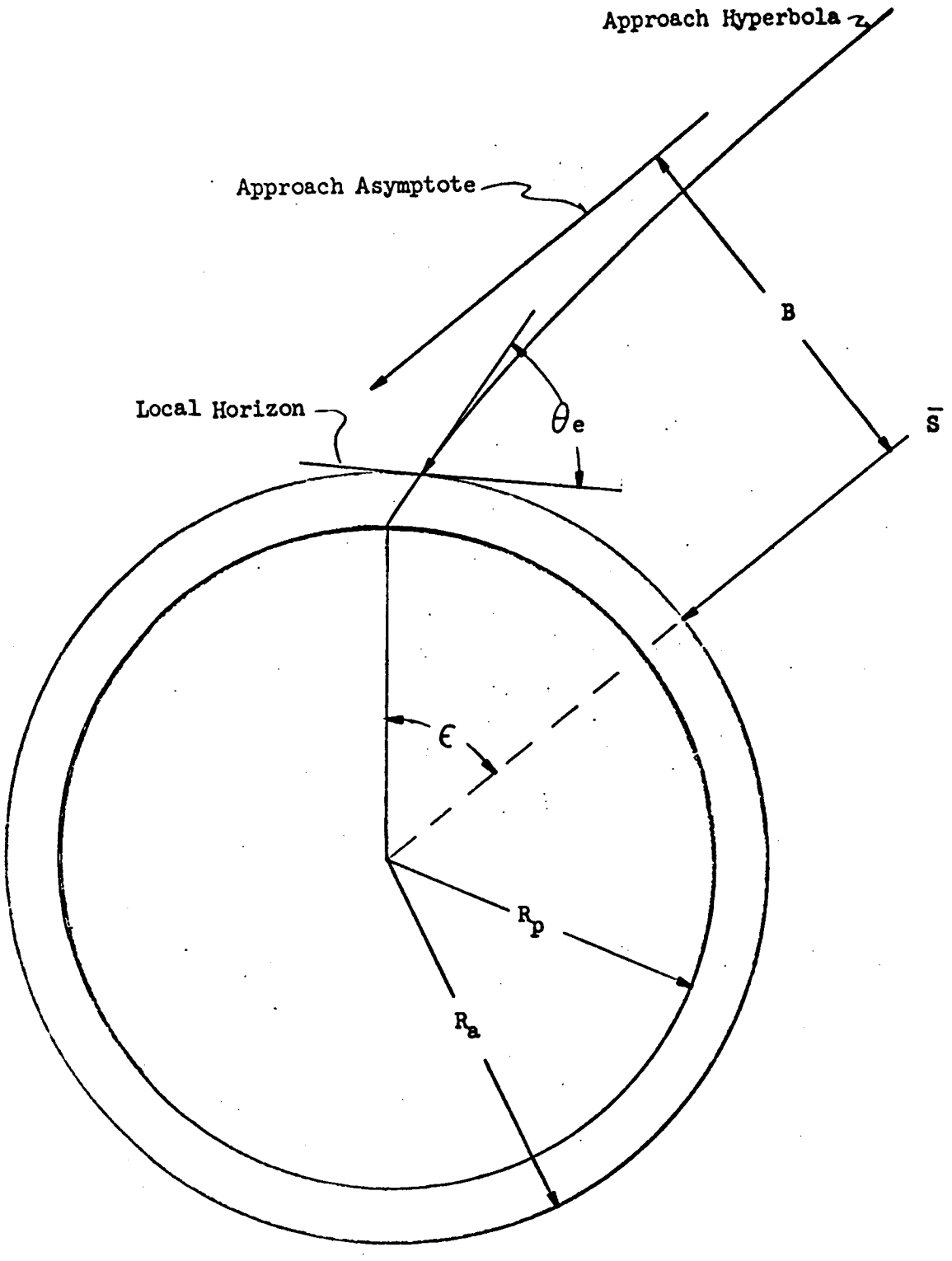


Figure 4.11. Bending Angle Versus Miss Parameter



R_p = 3378 km

R_a = 3621 km

Figure 4.12. Range Angle, ϵ, and Entry Angle θ_e

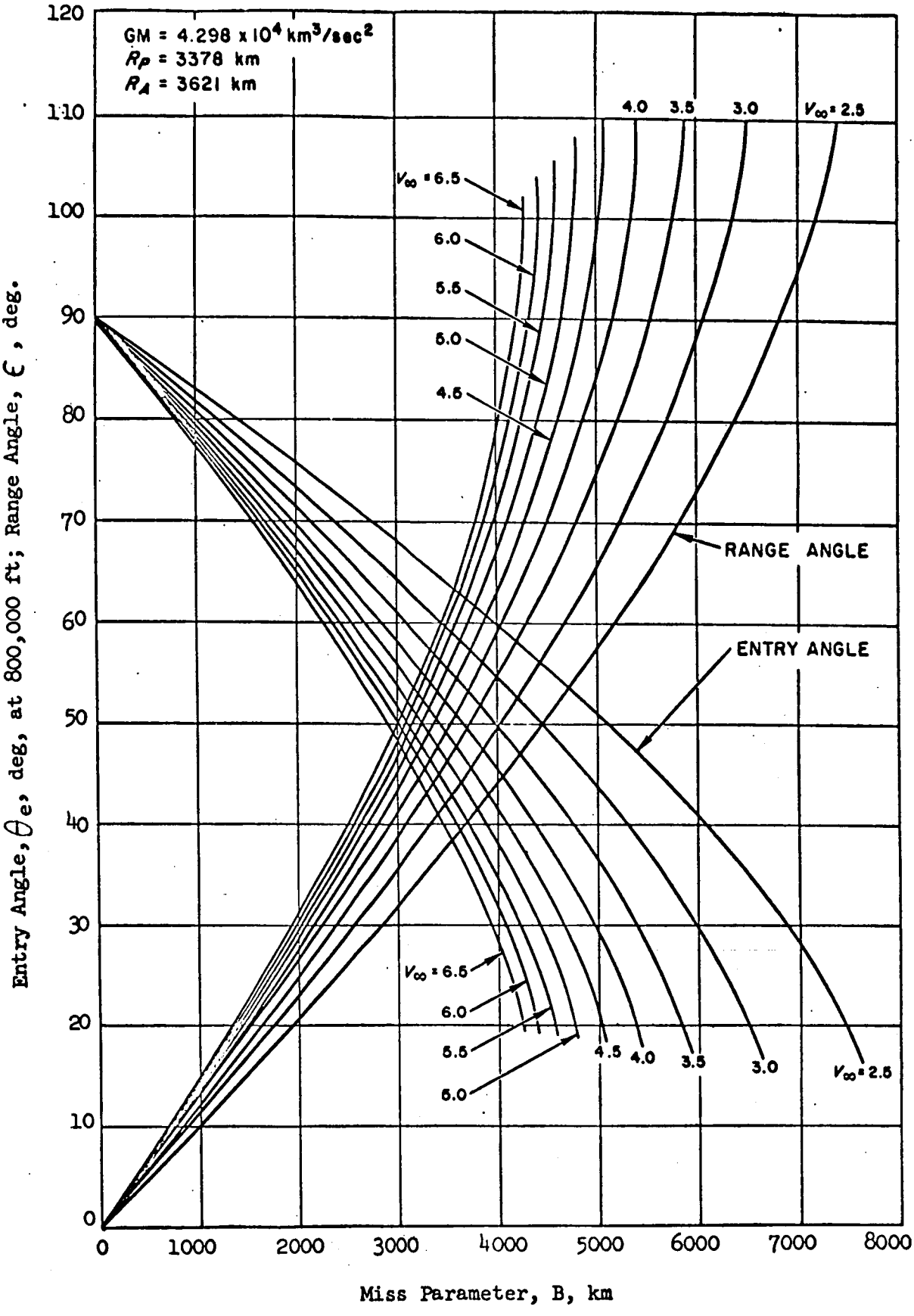


Figure 4.13. Entry Angle & Range Angle Versus Miss Parameter

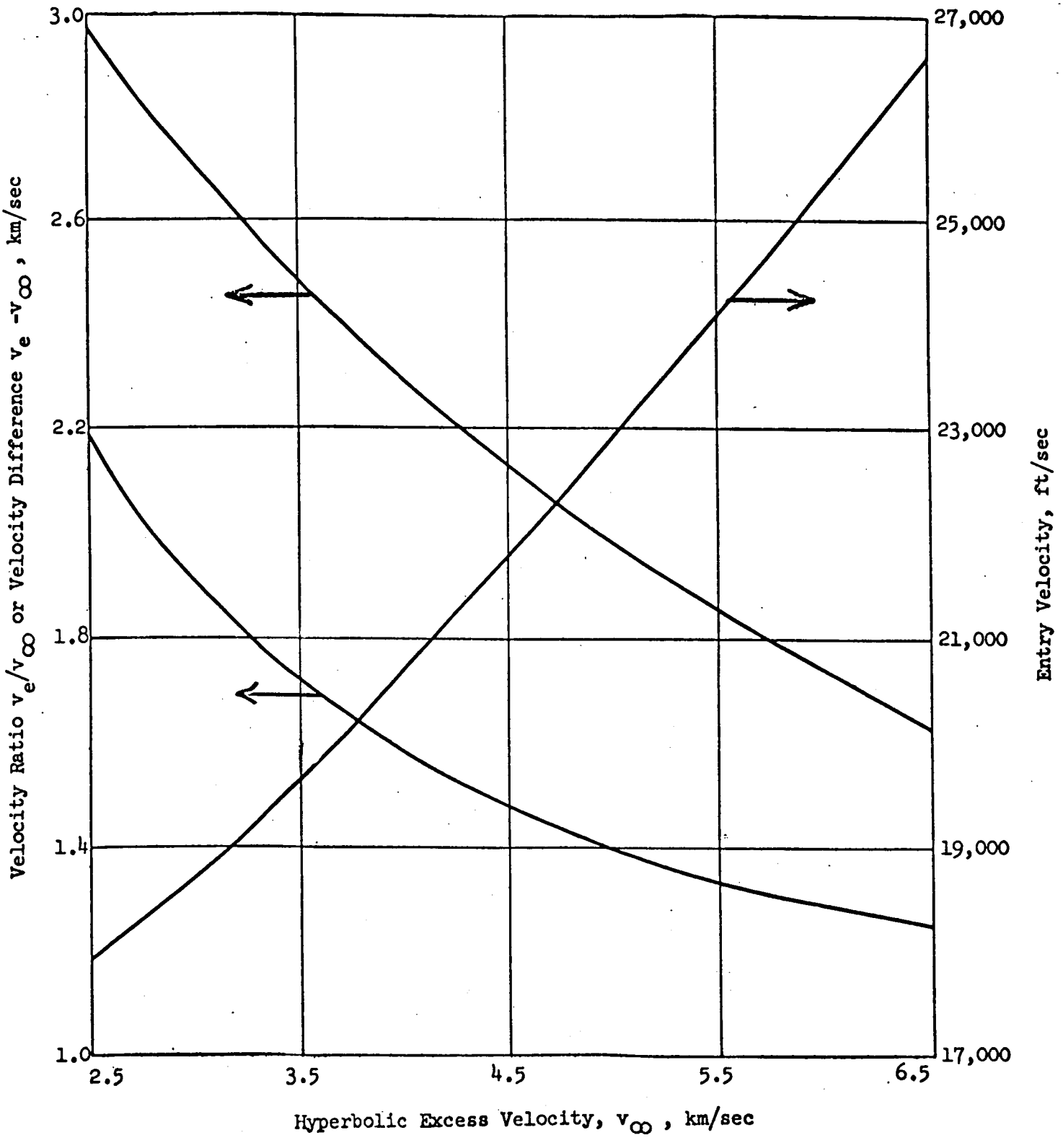


Figure 4.14. Entry Velocity Versus Hyperbolic Excess Velocity

It is possible now to consider the problems associated with separating a capsule from a spacecraft in the vicinity of the planet. One could consider a mission in which a small entry capsule was carried by a spacecraft to Mars then separated from the spacecraft and placed on an impact trajectory to the planet. The spacecraft could then serve as a relay station between the capsule and the Earth, or the capsule could transmit the information directly back to Earth. In either or both cases, it may be desirable for the spacecraft to perform some experiments, such as television, when it flys by the planet. Thus we see that there exists some rather important geometrical relations between the planet, the spacecraft, the capsule and the Earth.

First, let us look at the magnitude and direction of the maneuver required to place the capsule on an impact trajectory. If the spacecraft is targeted at some aiming point in the R - T plane with a value $\bar{B}_s$ and it is desirable to have the capsule targeted to an aiming point in the R - T plane with a value $\bar{B}_c$, then there exists a deflection distance, $\bar{D}$ as shown in Figure 4.15.

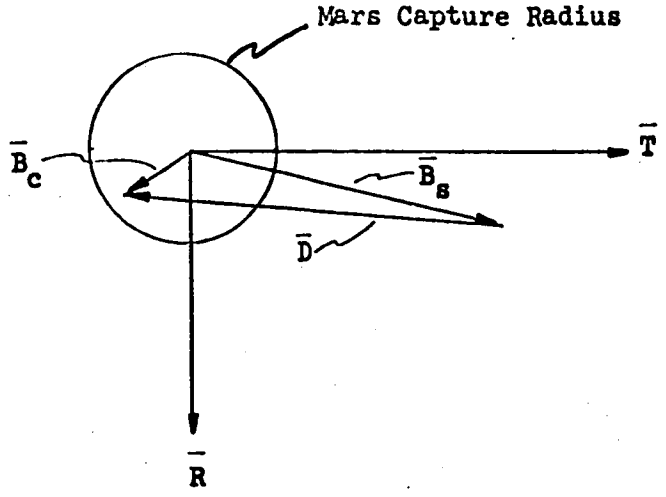


Figure 4.15. Spacecraft - Capsule Aim Diagram

From Figure 4.15 it can be seen that $\bar{D} = \bar{B}_c - \bar{B}_s$. The spacecraft approach asymptote $\bar{S}_s$ and $\bar{D}$ form a plane. This plane is perpendicular to the R - T plane and shall be called the maneuver plane. Now using the concept of this maneuver plane, it is possible to determine the magnitude and direction of the required capsule deflection maneuver. Another important parameter is the difference in arrival time at the planet between the capsule and the spacecraft. For example, if the capsule is transmitting to the spacecraft, then it is important for the capsule to arrive at the planet before the spacecraft. This difference in arrival time, ΔT_a, places some important considerations on the mission design as will be seen later.

The time before the spacecraft encounter at which the maneuver is made is T_f. With these parameters, a good approximation to the actual geometry can be made assuming the motions to be rectilinear. This is shown in Figure 4.16 where the plane of the paper corresponds to the maneuver plane. The communication range at encounter, X, is given by

$$X = \sqrt{D^2 + (v_\infty \Delta T_a)^2}$$

where $v_\infty \Delta T_a$ is the distance the spacecraft moves after the capsule impacts. These three parameters, X, D and $v_\infty \Delta T_a$, uniquely determine the desired encounter geometry. The velocity increment, Δv, applied to the capsule has two components

$v_1 = \Delta v \cos$ and

$v_2 = \Delta v \sin$ such that

$v_1 T_f = v_\infty \Delta T_a$ and

$v_2 T_f = D$ and since

$$\tan \alpha = \frac{v_2}{v_1} \quad \text{then}$$

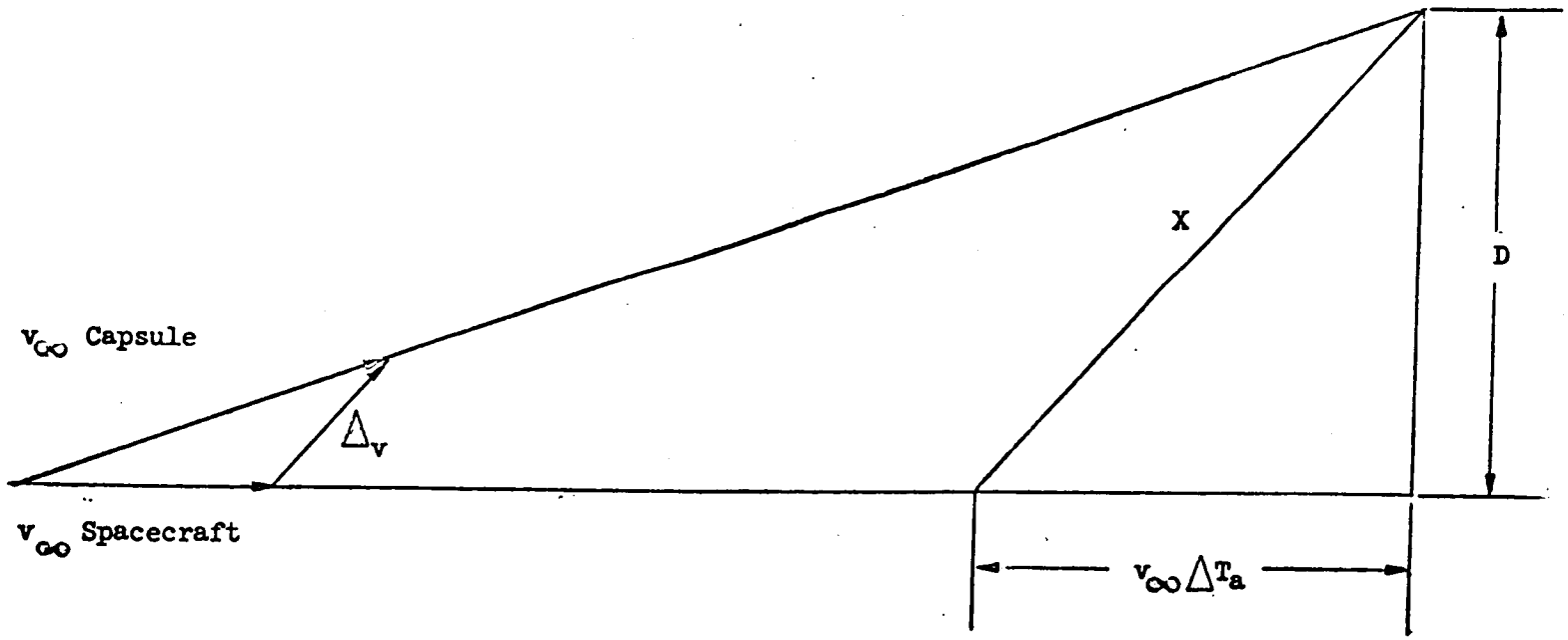


Figure 4.16. Maneuver Plane Geometry

$$\tan \alpha = \frac{D}{v_\infty \Delta T_a} \qquad \text{therefore}$$

$$\Delta v \, T_f = X \qquad \text{or}$$

$$\Delta v = \frac{X}{T_f} \, .$$

Thus, the required velocity increment is seen to vary inversely with the separation time and directly with the communication distance, X, at encounter, and the application angle, α, is given by

$$\alpha = \tan^{-1} \left(\frac{D}{v_\infty \Delta T_a} \right)$$

In referring to Figure 4.16, it is important to realize that the actual geometry is significantly different since the v_∞ of the capsule and spacecraft are almost parallel and have essentially the same magnitude.

Some first order approximations to the accuracy of such a maneuver can be made as follows: There are two main error sources to be considered, σ_v, the error in the total velocity increment, and σ_p the error in pointing (or direction), about two orthogonal axis. The velocity error will be assumed to be a fixed percentage of the total magnitude and the error in pointing an absolute value. These two errors map into an aiming point error with three orthogonal components: two in the aiming plane and one normal to the aiming plane. The two "in plane" errors have directions along $\bar{S}$ and $\bar{D}$; the "out of plane" error is normal to $\bar{S}$ and $\bar{D}$. These three errors will be defined as follows:

σ_{in} - the in plane error in the D direction,

σ_s - the in plane error in the S direction,

σ_{out} - the out of plane error

The value of σ_{in} is given by

$$\sigma_{in} = \sqrt{(\sigma_v D)^2 + (\sigma_p v_\infty \Delta T_a)^2}$$

This can be seen in Figures 4.17 and 4.18 and the following deviation.

Both the pointing and velocity errors map into position errors as shown in Figures 4.17 and 4.18. From similar triangles it can be seen that

$$\frac{\sigma_v X}{X} = \frac{\sigma_1}{D} \quad \text{or}$$

$$\sigma_1 = \sigma_v D \quad \text{and}$$

$$\frac{\sigma_p X}{X} = \frac{\sigma_2}{v_\infty \Delta T_a} \quad \text{or}$$

$$\sigma_2 = \sigma_p v_\infty \Delta T_a$$

Now the total error in the D direction is the RSS of these two or

$$\sigma_{in} = \sqrt{\sigma_1^2 + \sigma_2^2}$$

$$\sigma_{in} = \sqrt{(\sigma_v D)^2 + (\sigma_p v_\infty \Delta T_a)^2}$$

Similarly the inplane error in the $\bar{S}$ direction is

$$\sigma_s = (\sigma_p D)^2 + (\sigma_v v_\infty \Delta T_a)^2$$

$$\frac{\sigma_p X}{X} = \frac{\sigma_3}{D} \quad \text{or}$$

$$\sigma_3 = \sigma_p D \quad \text{and}$$

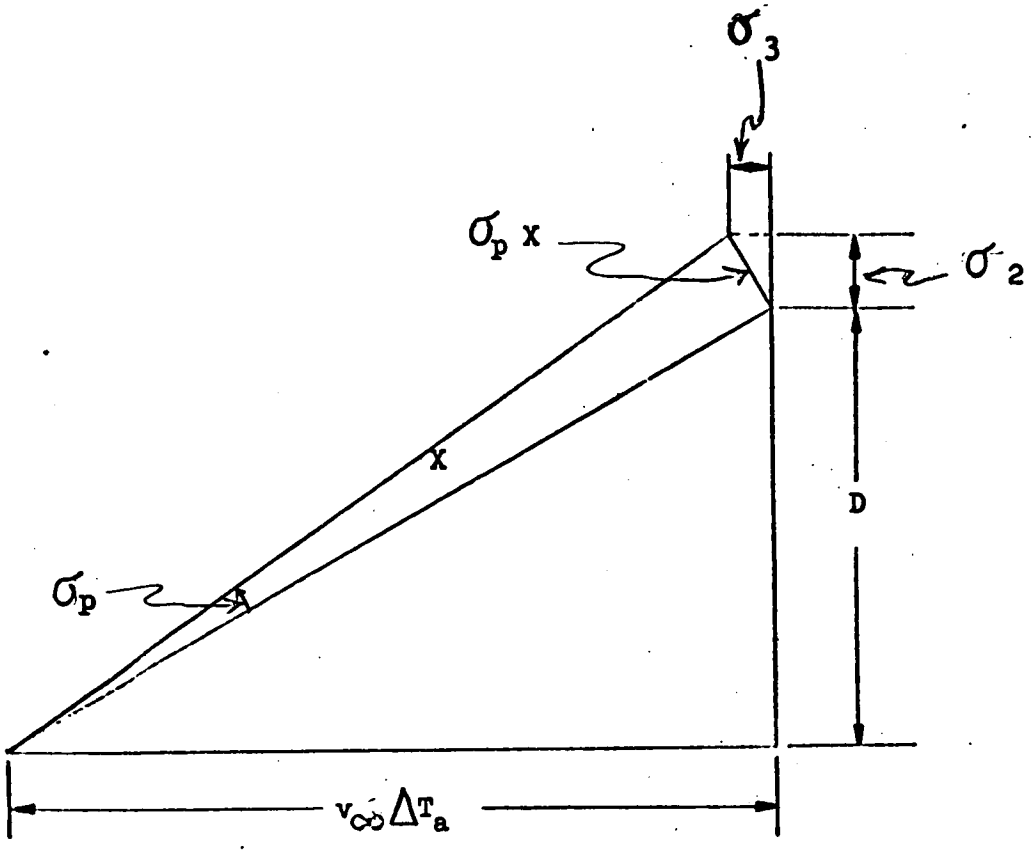


Figure 4.17. Pointing Error Effects

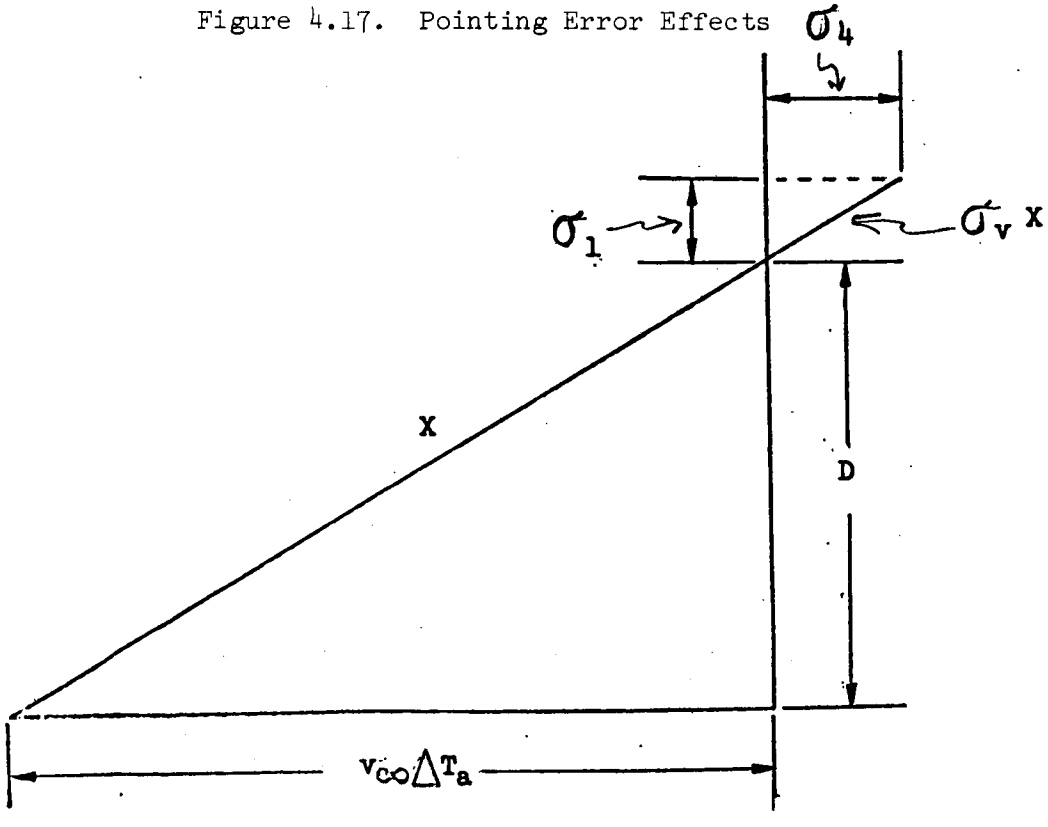


Figure 4.18. Velocity Error Effects

$$\frac{\sigma_v X}{X} = \frac{\sigma_4}{v_\infty \Delta T_a} \quad \text{or}$$

$$\sigma_4 = \sigma_v \, v_\infty \, \Delta T_a \quad \text{and then}$$

$$\sigma_s = \sqrt{\sigma_3{}^2 + \sigma_4{}^2} \quad \text{or}$$

$$\sigma_s = \sqrt{(\sigma_p D)^2 + (\sigma_v v_\infty \Delta T_a)^2}$$

And finally the component out of the plane is simply

$$\sigma_{out} = \sigma_p X$$

The in plane error in the S direction is important in determining the error in arrival time, σ_t, where

$$\sigma_t = \frac{\sigma_s}{v_\infty}$$

It should be noted that these errors are only the relative errors between the spacecraft and the capsule. In addition to these, there are orbit determination errors which account for the uncertainty of where the fly by trajectory is with respect to the actual position of the planet. This error must be added (RSS) to these errors. The σ_{in} and σ_{out} then map in the R - T plane as shown in Figure 4.19.

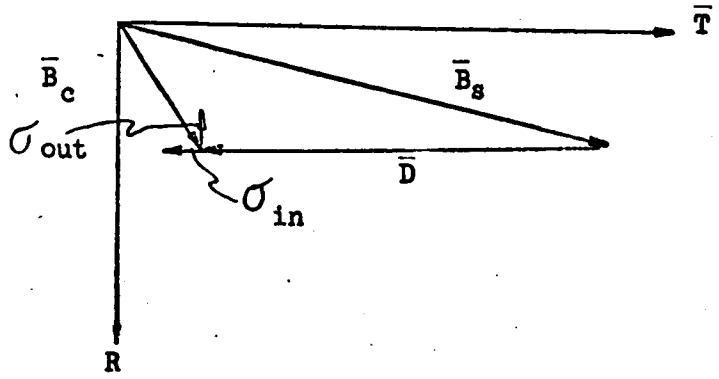
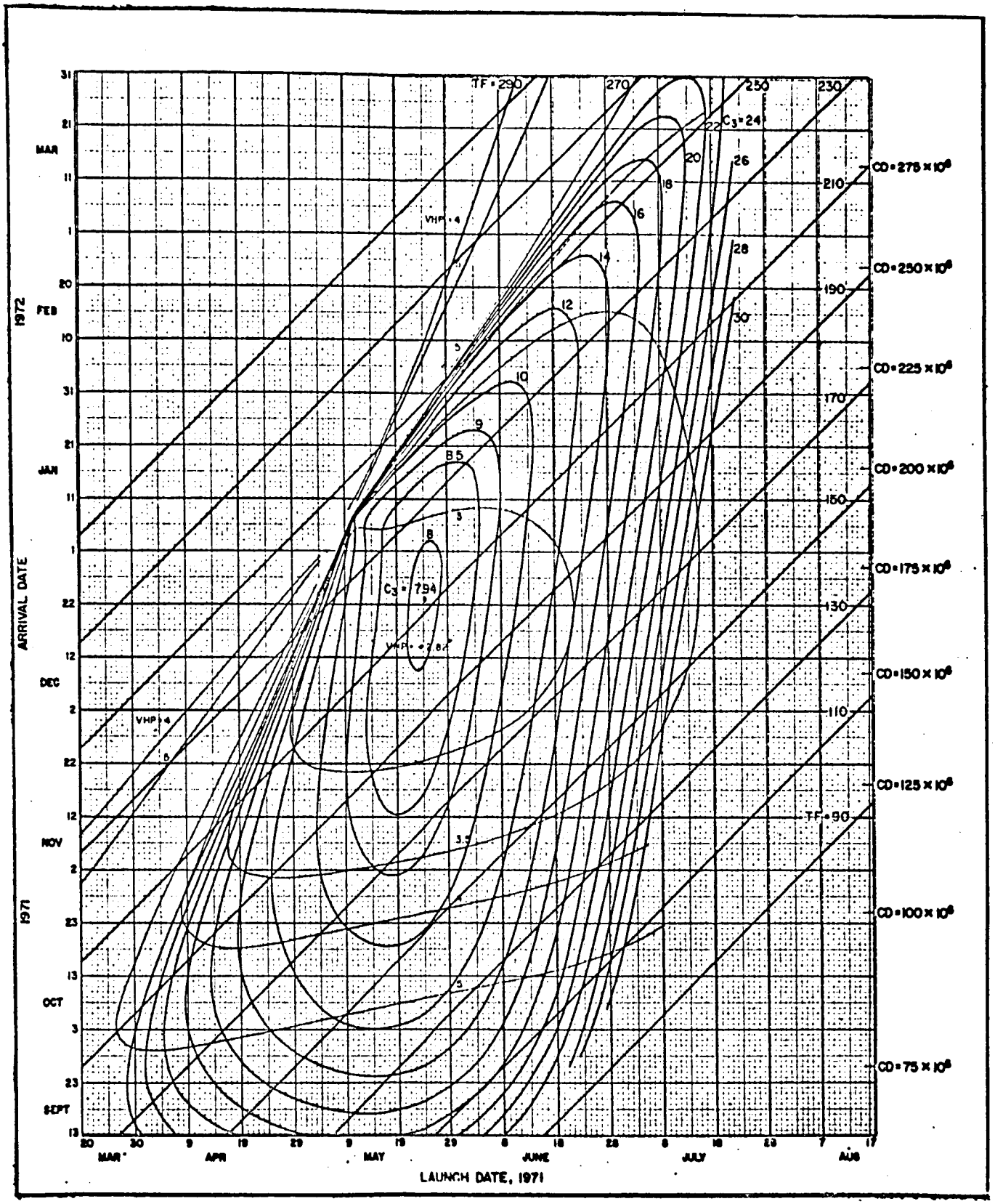


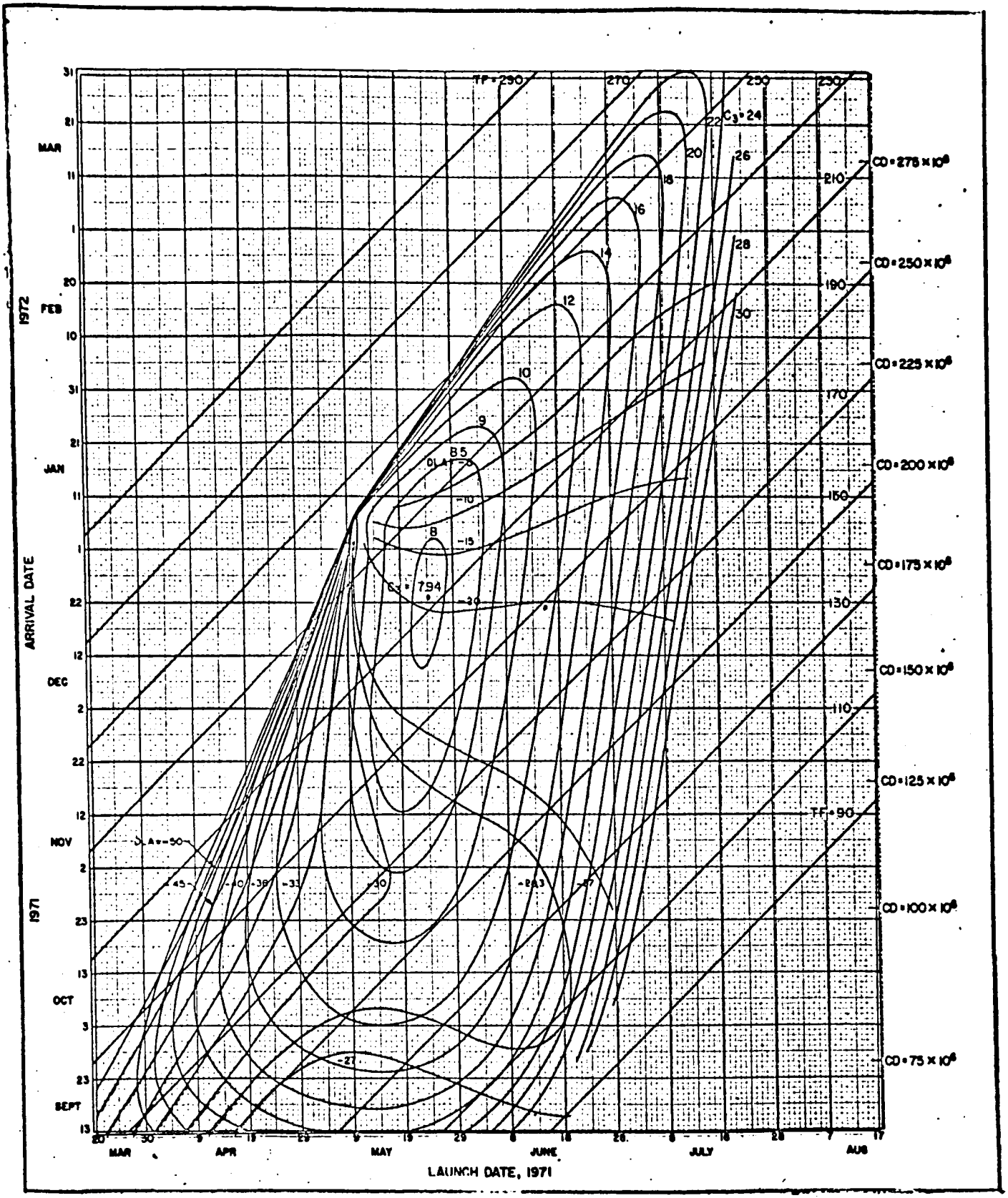
Figure 4.19. In Plane and Out of Plane Excitation Errors

Figures 4.20 through 4.26 are basic design charts for the 1971 Type 1 trajectories. The closed containers are values of C_3, which can be directly related to the payload capability of any launch vehicle.



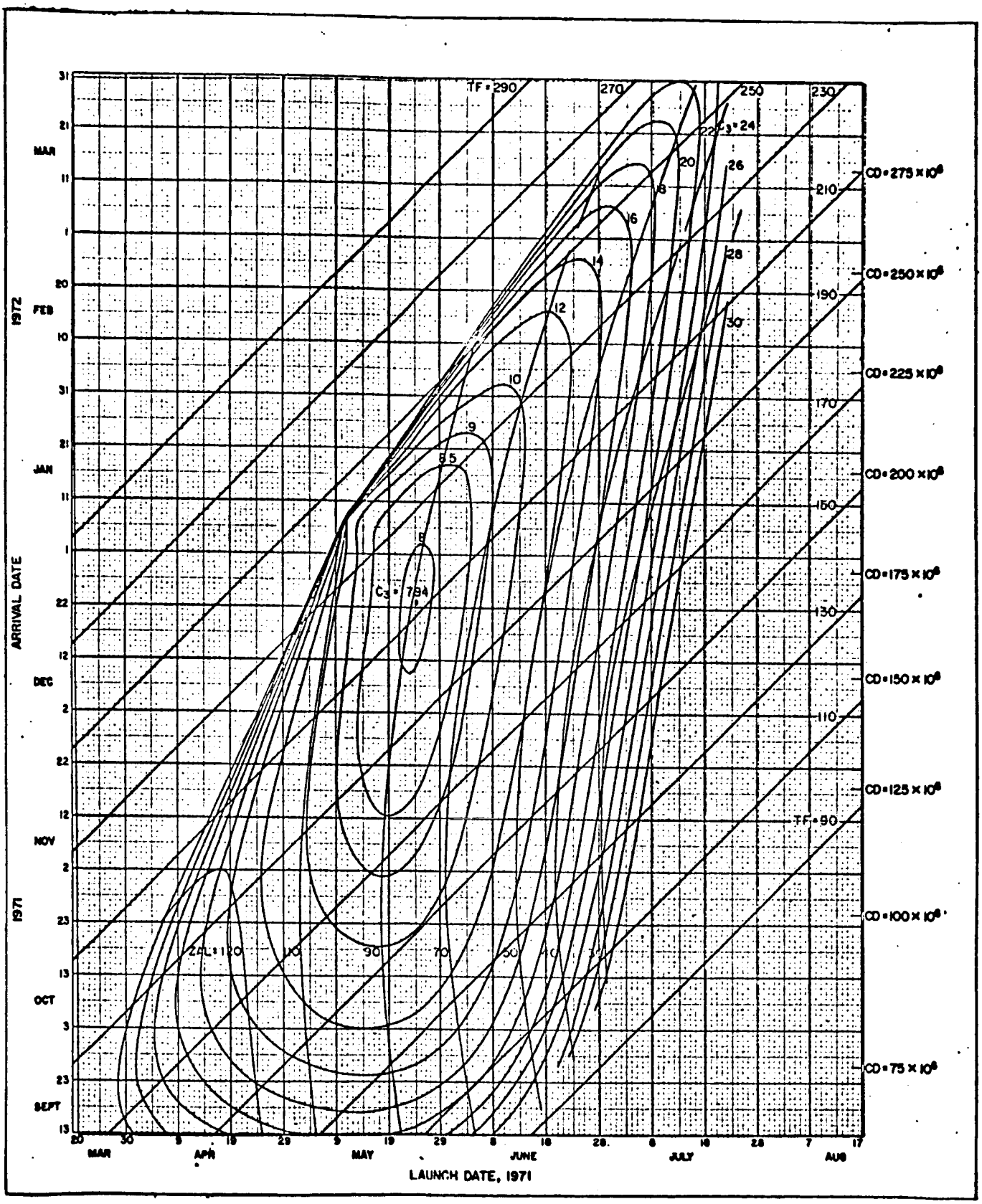
Basic Trajectory Design Chart, 1971, Type 1

Figure 4.20. Hyperbolic Excess Velocity Relative to Mars



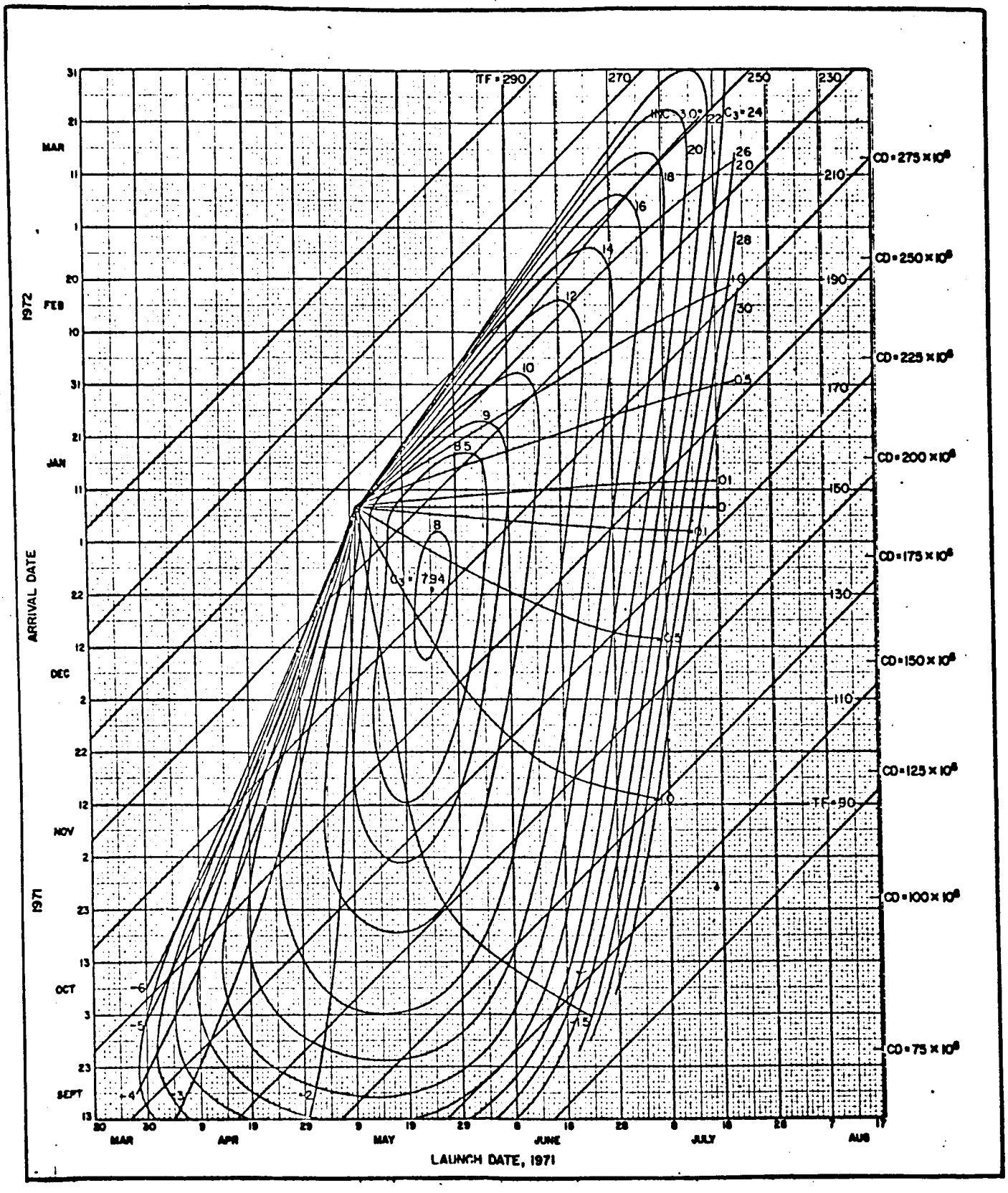
Basic Trajectory Design Chart, 1971, Type 1

Figure 4.21. Declination of the Geocentric Departure Asymptote



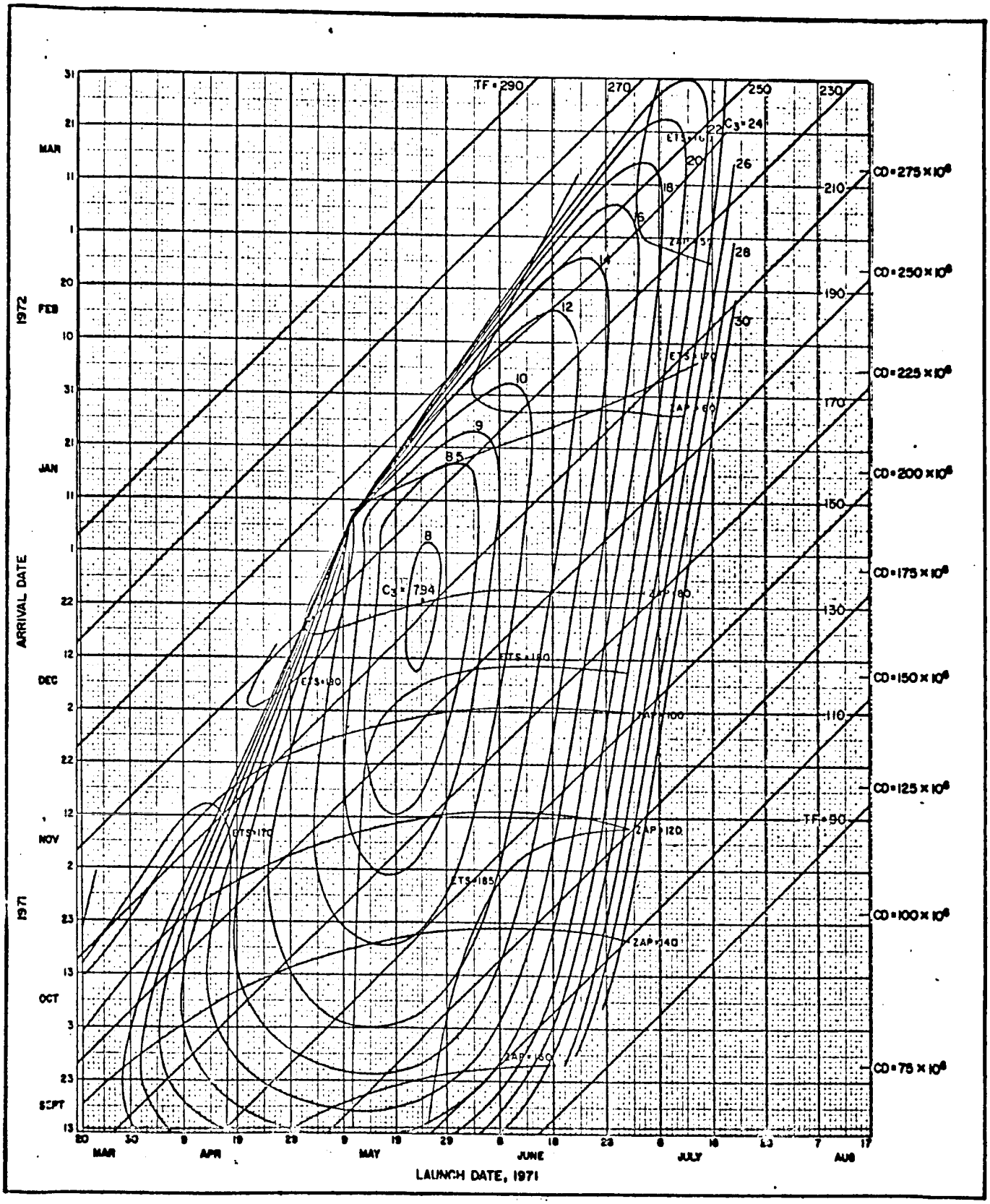
Basic Trajectory Design Chart, 1971, Type 1

Figure 4.22. Angle Between the Sun-Earth Vector and Departure Asymptote



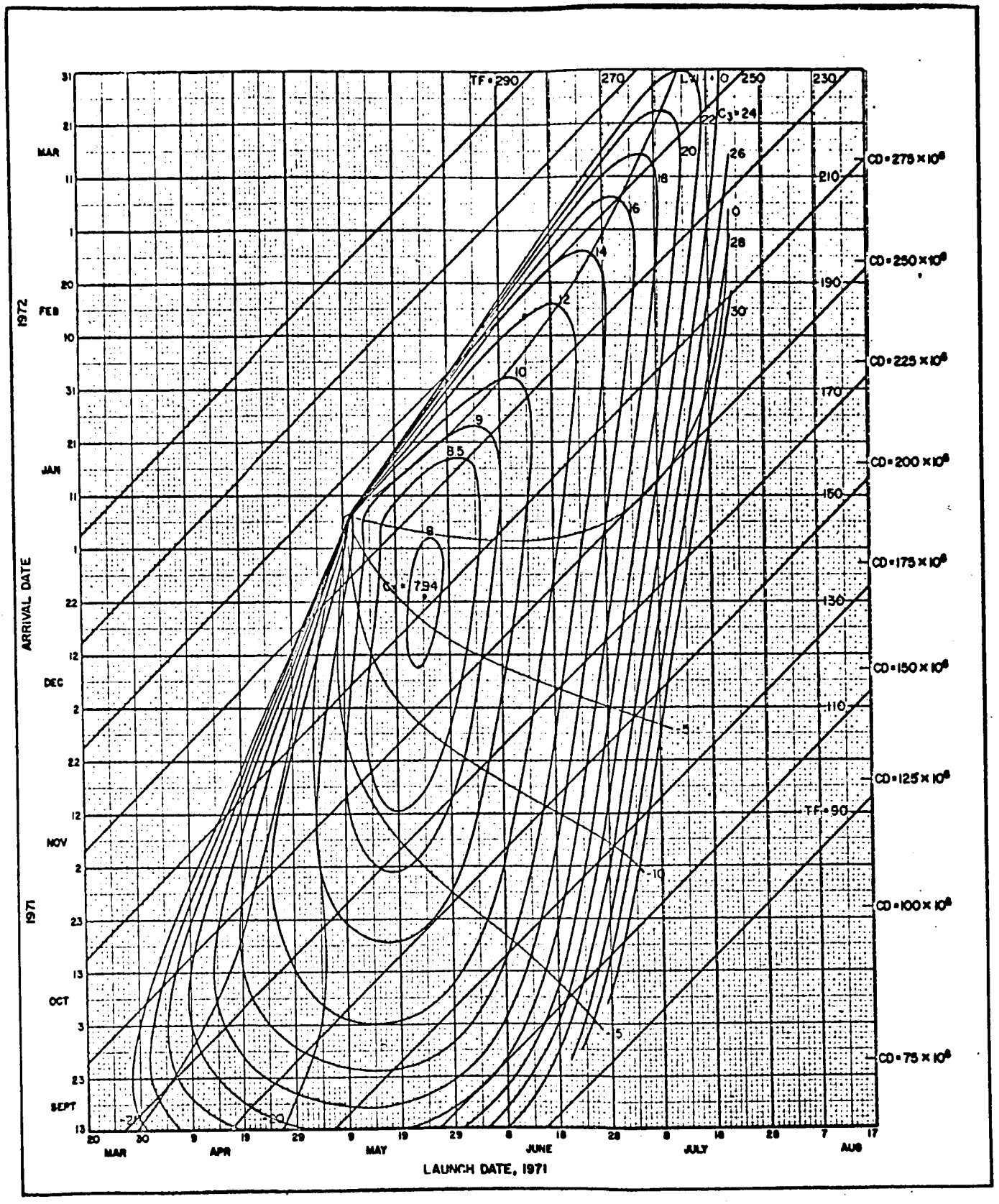
Basic Trajectory Design Chart, 1971, Type 1

Figure 4.23. Inclination of the Transfer Plane to the Ecliptic Plane



Basic Trajectory Design Chart, 1971, Type 1

Figure 4.24. ZAP and ETS



Basic Trajectory Design Chart, 1971, Type 1

Figure 4.25. Latitude of the Vertical Impact Point on Mars

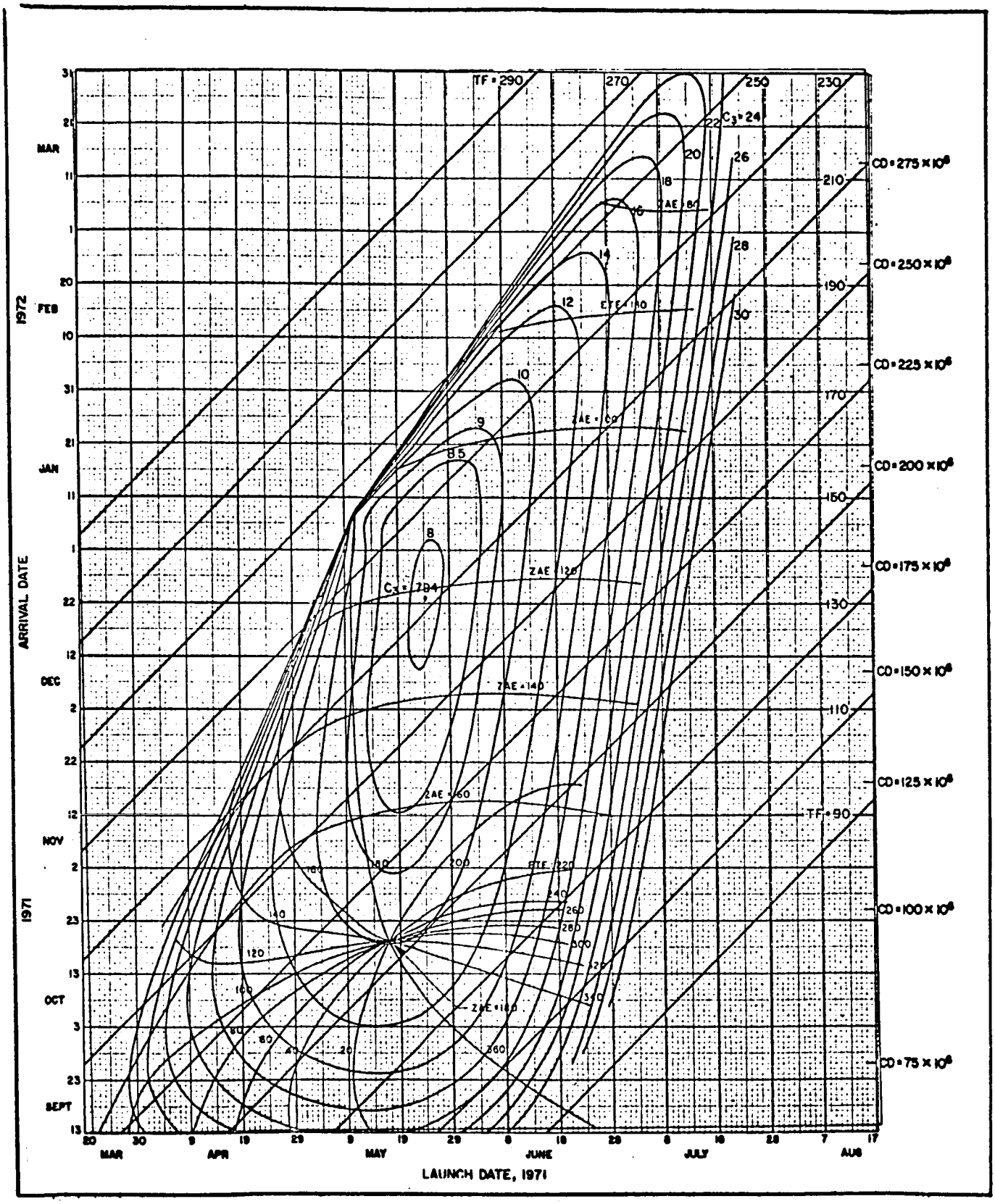


Figure 4.26. ZAE and ETE

IV SPACE COMMUNICATIONS (J. STIFFLER)

1 INTRODUCTION

A space mission is no more successful than its telemetry. For it is only through the telemetry received from the vehicle that we are able to determine how it is performing, what the engineering and scientific measurements are yielding, and, with the possible exception of large satellites in near-earth orbits, whether the vehicle is even there or not.

Wireless telemetry has been a reality for about seventy years now, and has been used extensively for forty or fifty years. In that time radio, television, and radar techniques have been developed to extremely high levels of performance. Commercial radio and television has become a way of life; commercial aviation depends upon radio and radar, and even amateurs operate their own radio transmitters. Why, then, should there be any problems associated with telemetry in the space age? It would seem that all the efforts should be devoted to the less tried engineering pursuits, such as rocket design. Telemetry, albeit important, is well understood.

A little reflection, however, suggests a number of significant differences between space telemetry and surface telemetry. The most striking of these lies in the distances involved. Clearly, until recently, no attempt was ever made to transmit information more than about 12,000 miles, since no two points on earth are separated by more than that distance. (Even communication at distances of more than a few thousand miles depended upon rather unpredictable meterological phenomena and could not be relied upon.) Yet the nearest

neighbor to the earth is about twenty times farther away than that and the planet nearest to the earth more than 2000 times as distant. Since the power at the receiver is inversely proportional to the square of the distance between it and the transmitter, the power received from a transmitter on Venus would be only $\frac{1}{4,000,000}$ th as great as the power received from the same transmitter placed on the opposite side of the earth.

The second major difference between space telemetry and surface telemetry rests in the constraints placed on the transmitter and receiver in the space vehicle. In contrast to commercial radio and television, in which the transmitter can be large and complex whereas the receivers must be kept small and inexpensive, the transmitter in the spacecraft-to-ground telemetry link is limited by weight and reliability constraints while the receiver, on the ground, is relatively unconstrained. It is clearly not possible, or at least not practical, to include, as part of a spacecraft, a 50,000 watt transmitter complete with a steam turbine to generate the power and an 80 foot antenna to transmit it. The equipment must be kept as small and as reliable as possible, and must be fully automatic; the power required can be no more than that supplied by the generators and batteries on board which, in turn, must be kept within reasonable weight limitations. Thus, in practice, the transmitted power is limited to a few watts.

The receiver, however, may reasonably involve large antennas, a crew of operators, and even a digital computer. The ground-to-vehicle link is more conventional in its constraints, although the reliability demands and the operating conditions to which the receiver is subject

are, of course, considerably more severe than in the usual situation.

Granting, then, that space telemetry does, indeed, pose new and unique problems, what are the means which have been adopted for a solution? They, in general, fall into one of two categories: 1) improved components, and 2) improved data handling and modulation systems.

The discussion of these techniques is the purpose of this chapter. We begin by reviewing some mathematical techniques and introducing some fundamental concepts in Section one. In Section Two we investigate some of the methods whereby the effective signal power at the receiver can be vastly increased through improved component design. We then discuss, in some detail in Section Three, the somewhat conventional modulation techniques and proceed to investigate the more recent pulse modulation schemes in Section Four. Having observed some definite advantages inherent in pulse modulation we then discuss, in Section Five, the data handling efficiencies possible when working with pulsed, or sampled, data. Section Six is concerned with the related problems of ranging and, briefly, telemetry synchronization. And, finally, Section Seven discusses the telemetry systems which have actually been used in the Ranger and Mariner programs and some of the recent innovations used in the earth-based receiving equipment.

1.1 Fundamentals

In order to discuss, in any detail, the essentials of modern communications it is necessary first to define some terms and to review some mathematical concepts. It is hoped that this chapter will help to provide a heuristic clarification of some of the concepts such as bandwidth and noise so fundamental to the understanding of telecommunication techniques.

1.1.1 Fourier Series

Let $f(t)$ be a function periodic in time with a period T such that $f(t) = f(t + nT)$ $n = 0, \pm 1, \pm 2, \ldots$ and with the property that

$$\int_{-T/2}^{T/2} |f(t)|\, dt < \infty$$

Then subject to some rather general conditions $f(t)$ can be expressed as a Fourier series:

$$f(t) = \frac{a_o}{T} + \frac{2}{T} \sum_{n=1}^{\infty} (a_n \cos \omega_n t + b_n \sin \omega_n t) \qquad (1.1)$$

where $\omega_n = \frac{2\pi n}{T}$. Thus f(t) may be considered to be a weighted sum of sinusoids of frequencies $f_n = \frac{\omega_n}{2\pi} = \frac{n}{T}$. The Fourier coefficient a_o may be evaluated by integrating both sides of equation (1.1) over one period

$$\int_{-T/2}^{T/2} f(t)\, dt = a_o \qquad (1.2a)$$

Similarly a_n and b_n can be evaluated by multiplying both sides of equation 1 by $\cos \omega_n t$ and $\sin \omega_n t$ respectively and integrating the product over one period:

$$a_n = \int_{-T/2}^{T/2} f(t) \cos \omega_n t\, dt$$

$$b_n = \int_{-T/2}^{T/2} f(t) \sin \omega_n t\, dt \qquad (1.2b)$$

While we have equated the function f(t) to its Fourier series, it should be noted that this equality is valid only at the points of continuity of f(t).

As an example consider the function illustrated in figure 1.1. The Fourier series expansion of the periodic function f(t) is easily determined:

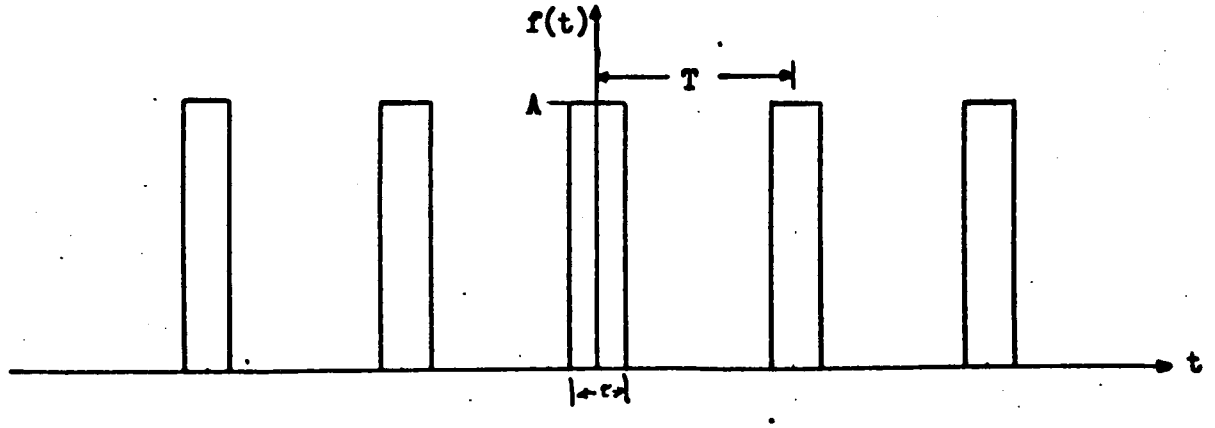


Figure 1.1. A Periodic Function

$$f(t) = \frac{A\tau}{T} + 2 \sum_{n=1}^{\infty} \frac{a_n}{T} \cos \omega_n t \qquad (1.3)$$

where $a_0 = A\tau$

and $$a_n = A \int_{-\tau/2}^{\tau/2} \cos \omega_n t \, dt$$

$$= 2A \frac{\sin \omega_n \tau/2}{\omega_n}$$

1.1.2 Power Spectra

Let y(t) be a periodic function

$$y(t) = \frac{a_0}{T} + \frac{2}{T} \sum_{n=1}^{\infty} (a_n \cos \omega_n t + b_n \sin \omega_n t) \quad \omega_n = \frac{2\pi n}{T} \qquad (1.4)$$

The average power in y(t) is defined as

$$P_{ave} = \frac{1}{T} \int_{-T/2}^{T/2} y^2(t) \, dt \; . \qquad (1.5)$$

(if y(t) is a voltage level, for example, P_{ave} represents the average power dissipated by passing y(t) through a one ohm resistor.) If P_{ave} is finite, we can substitute equation (1.4) into equation (1.5) and obtain

$$P_{ave} = \frac{a_0^2}{T^2} + \frac{4a_0}{T^3} \int_{-T/2}^{T/2} \sum_{n=1}^{\infty} (a_n \cos \omega_n t + b_n \sin \omega_n t) dt$$

$$+ \frac{4}{T^3} \int_{-T/2}^{T/2} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} (a_n \cos \omega_n t + b_n \sin \omega_n t)(a_m \cos \omega_n t + b_m \sin \omega_m t) dt$$

The first integral on the right is identically zero, while the second integral can be written as

$$\frac{2}{T^3}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left\{\int_{-T/2}^{T/2}(a_n a_m + b_n b_m)\cos(\omega_n-\omega_m)t\,dt\right.$$

$$+\int_{-T/2}^{T/2}(a_n a_m - b_n b_m)\cos(\omega_n+\omega_m)t\,dt$$

$$+\int_{-T/2}^{T/2}(b_n a_m - a_n b_m)\sin(\omega_n-\omega_m)t\,dt$$

$$\left.+\int_{-T/2}^{T/2}(a_n b_m + b_n a_m)\sin(\omega_n+\omega_m)t\,dt\right\}$$

$$=\frac{2}{T^3}\sum_{n=1}^{\infty}\int_{-T/2}^{T/2}(a_n^2+b_n^2)dt=\frac{2}{T^2}(a_n^2+b_n^2)$$

Thus,

$$P_{ave}=\left(\frac{a_0}{T}\right)^2+\frac{2}{T^2}(a_n^2+b_n^2) \qquad (1.6)$$

But $\left(\frac{a_0}{T}\right)^2$ is just the power in the constant (d.c.) term of equation (1.1) while $\frac{2}{T^2}(a_n^2+b_n^2)$ represents the average power at the n^{th} frequency component. Thus, the average power of a periodic time function is just the sum of the average powers at each of its frequency components. This result is known as Parseval's theorem.

The power spectrum $\phi(\omega)$ of the function f(t) is defined as the average amount of power at each of the frequency components $f_n = n/T$. It is typically illustrated as in figure 1.2, the height

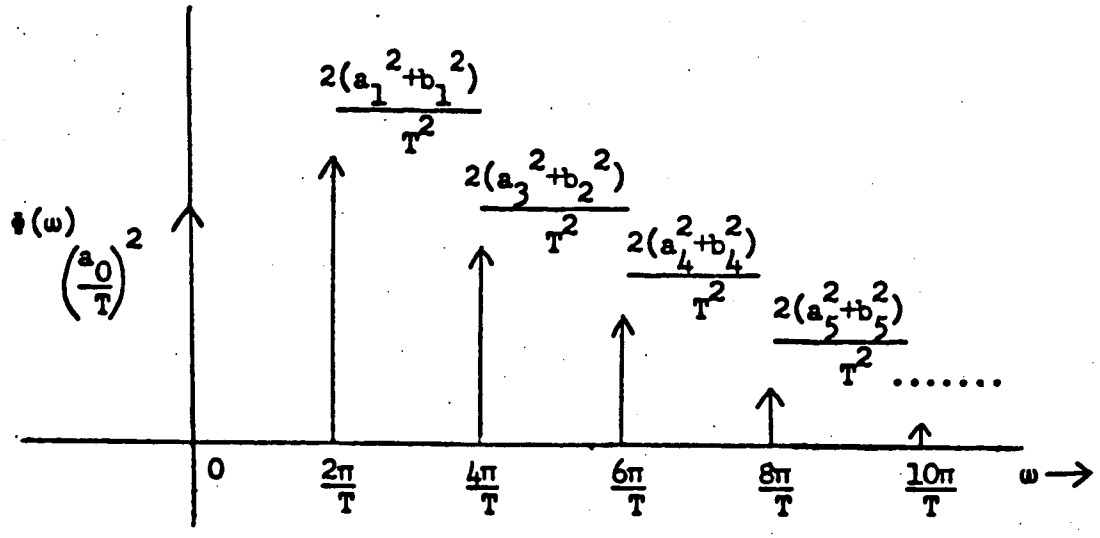


Figure 1.2. A Discrete Power Spectrum

of the arrows being proportional to the amount of power at that frequency.

1.1.3 Generalization to Non-Periodic Time-Functions and Continuous Spectra

To a communication theorist a periodic function is a pretty dull thing (once you've seen one period you've seen 'em all). The very nature of communication demands that the information transmitted is unexpected. If it is purely repetitive, why transmit it? (We shall actually encounter some reasons, shortly.) Usually, however, the transmitted signal will not be periodic and the techniques we have been discussing are not directly applicable. They can, however, be generalized in a straightforward way to include the signals typically encountered in a communications system.

Any detailed discussion of the more general time functions which are of concern to a communications engineer is beyond the scope of these notes. A few general comments are in order, however, and will prove to be useful in the subsequent discussion. As we have noted, a time function is interesting from a communications point of view to the

extent that it is unpredictable. Observing the received signal $y(t)$ at time t_1 may provide only limited information regarding the value of that function at time t_2. The quantities $y(t_1)$, $y(t_2)$, etc., can be regarded as random variables; the associated time function $y(t)$ is denoted a <u>random process</u>.

Random processes can also be partially characterized by their power spectra. The power spectrum of a random process is fundamentally different from that of a periodic deterministic time function, however, in that it will no longer consist entirely of discrete components. The power spectrum of a random process might have the form shown in figure 1.3. Now, rather than the amount of power at a particular frequency $f = f_n$,

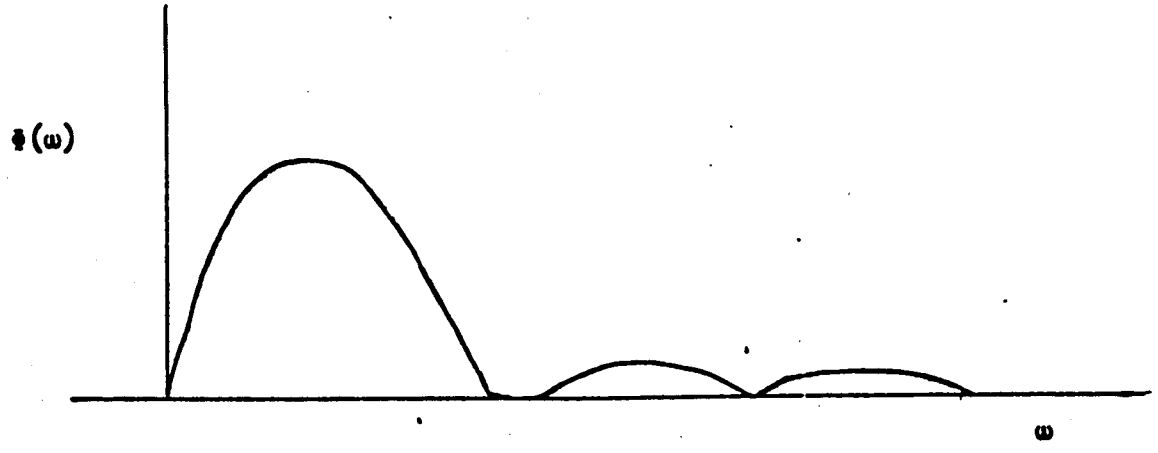


Figure 1.3. A Continuous Power Spectrum

the relevant quantity is the power exhibited in the frequency <u>interval</u> $f_0 < f < f_0 + \Delta f$. This is defined in terms of the quantity $\Phi(\omega)$ by the equation

$$P_{ave}\ (f_0 < f < f_0 + \Delta f) = \int_{f_0}^{f_0 + \Delta f} \Phi(2\pi f)\, df. \qquad (1.7)$$

The function $\Phi(\omega)$ is therefore called the <u>power spectral density</u>; its integral yields the average power of the process in question in any given frequency interval. The total average power of the

process is given by the integral of $\phi(\omega)$ over the entire frequency range:

$$P_{ave} = \int_0^\infty \phi(2\pi f)df \qquad (1.8)$$

Often, random processes exhibit some periodic behavior also, in which case the associated power spectrum will contain both continuous and discrete parts, as shown in figure 1.4.

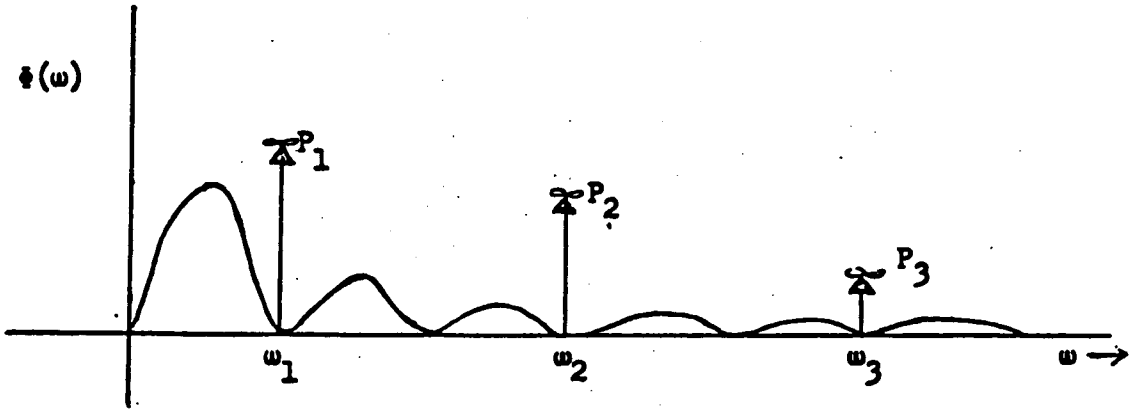


Figure 1.4. A Power Spectrum Having Both Continuous and Discrete Parts

The arrows again indicate that the process contains power at the discrete frequency components $f_1 = \omega_1/2\pi$, $f_2 = \omega_2/2\pi$, etc. The power spectral density at these points is infinite. For this reason, the arrows are often capped with infinity signs, the height of the arrows again generally being proportional to the amount of power at that frequency. The power at the discrete points of the spectrum is usually indicated explicitly as in figure 1.4. The total power is the process represented by figure 1.4 is given by

$$P_{ave} = \sum_{i=1}^{\infty} P_i + \int_0^\infty \phi(2\pi f)df.$$

1.1.4 Noise

Were it not for noise space communication would be relatively simple. The receiver antenna would be followed by an amplifier or amplifiers with enough gain to render the signal useful. While the receiver might be somewhat costly for very small signals, no signal would be too small and no transmitter would be too weak or too far removed for effective communication. Unfortunately, unpredictable random phenomena other than the signal is always present in any receiver. This noise is amplified by the same factor as the signal, and, while the voltage level may be amplified to more practical ranges, the signal is just as noisy, relatively, as it was before amplification. Noise must be counteracted by means other than amplification. In order to understand methods by which the effect of noise can be diminished it is necessary to consider for a moment the properties of the noise itself.

Any electricic system generates some random voltage or current fluctuations. A metallic resistor, for example, contains electrons which drift randomly from molecule to molecule. When this resistor is connected into a circuit the electron drift will produce a random current through the resistor and hence a voltage across its terminals. Such noise is called Johnson or thermal noise. It has been determined both experimentally and theoretically that for thermal noise, the noise power spectral density $\phi_n(\omega)$ is*

$$\phi_n(\omega) = k\,T \tag{1.9}$$

* The actual measured noise power of a resistor will depend upon the load into which it is operating. The value given here is that produced when the load is matched to the resistor. Since any mis-match affects the noise and signal in exactly the same way, it can have no bearing on the ratio of the two quantities; i.e., a mis-match is equivalent to an amplifier with a gain less than unity. Because of this it will be convenient here, and in the discussion to follow, to assume that the load impedence is always matched to the sources under consideration.

where k is Boltzmann's constant, $k = 1.38 \times 10^{-23}$ joule/°K, and T is the absolute temperature of the resistor in degrees Kelvin. While $\Phi_{in}(\omega)$ is not constant for all values of ω (this would indicate infinite power), the spectral density is flat out to extremely high frequencies where quantum effects begin to become significant.

Another commonly encountered noise source in an electronics system is the so-called shot noise. This is noise produced in vacuum tubes. Many millions of electrons are emitted randomly from the cathode of an electron tube each second. The average emission rate determines the average current. Nonetheless, because of the discrete properties of an electron, this current is not continuous but, rather, composed of many electron pulses. Thus, the instantaneous current fluctuates about this average. Since this fluctuation is random, and not produced by any input signal variations, it acts as noise. Shot noise may be accounted for by an equivalent noise source producing a noise power spectral density of $k T_o$ watts where k is Boltzman's constant, and T_o the effective noise temperature of the tube. This spectral density is flat out to frequencies on the order of the reciprocal of the time necessary for an electron to pass from the cathode to the anode. Although this frequency varies from tube to tube, it is typically on the order 10^9 cps.

Solid state devices, including transistors, also exhibit noise due to random electron fluctuations. Again this can be accounted for by including, in the circuitry of the transistor, an effective noise generator with a power spectral density of kT_t, where k is

again Boltzman's constant, T_t the effective noise temperature of the device. (Note that the effective noise temperature is not necessarily the actual temperature of the device in question.) The range of frequencies for which this spectral density remains cons-ant is generally somewhat less than that for vacuum tubes. However, it is usually safe to assume that the spectrum is flat over the frequency range for which the device is useful. Such noise is called white in analogy to the color white which represents all the visible spectrum.

In short, any electronic device generates noise, the spectral density of which is directly proportional to the product of its effective temperature (or to the temperature of one of its elements) and Boltzman's constant k. Any electronic system contains many of these independent noise generators.

All of these noise sources can generally be replaced theoretically by one noise generator at the input to the device which produces at the output a noise power equivalent to that that is actually generated by the combined action of the separate sources. The total noise spectral density needed at the input to account for the observed output is often written $N_O = K\ T_{eff}$ where T_{eff} is designated the effective noise temperature of the system in degrees Kelvin. Since the input signal and the effective input noise are equally amplified by the system, the signal-to-ncise ratio remains the same at the output as at the input. Thus, knowing the effective temperature of the system, it is only necessary to determine the input signal power in order to specify the output signal-to-noise ratio.

1.1.5 Bandwidth

The bandwidth of a signal is a measure of the width of its power spectrum in cycles per second. The bandwidth is of concern for several reasons. First, it is possible by using filters effectively to eliminate from the input to a receiver all frequencies outside a prescribed interval. Thus different communication systems can operate in the same vicinity without mutual interference as long as they are all restricted to operate over non-overlapping portions of the frequency spectrum. In commercial radio, the Federal Communications Commission prescribes the frequency internal within which every broadcaster must limit his transmitted signal. The radio receiver is tuned by changing the frequency range over which the input filter will pass a signal. In order to maximize the number of communication systems which can operate simultaneously, it is desirable to restrict, as much as possible, the bandwidth necessary to transmit the desired information. As we shall discover, however, other considerations will illustrate advantages to increasing rather than decreasing the bandwidth of the transmitted signal.

Another important reason for considering the bandwidth of a signal lies in the efficient design of the receiver. As we have observed, the noise spectrum tends to be very broad. If the desired signal were restricted to the frequency range $f_1 \leq f \leq f_2$, it would be disadvantageous to design a receiver which passed frequencies outside this range. To do so would only increase the noise power in the receiver without any increase in the signal power.

What is the relationship between a time function f(t) and its bandwidth? When f(t) is periodic we know that it can be expanded in terms of its Fourier series. Consider the example of figure 1.1. Its Fourier series expansion is that of equation 3. Since a_n is never zero for any finite frequency, f(t) requires an infinite bandwidth to be transmitted exactly. But note that the power at each frequency component, $\frac{2a_n^2}{T^2}$, does decrease rapidly as the frequency is increased. It can be verified, in fact, that about 95% of the total power is included in the interval $0 \leq f \leq \frac{1}{\tau}$. It would seem, then, that filtering out all frequencies greater than $f = \frac{1}{\tau}$ would alter the waveform but little.

This inverse relationship between time and frequency is quite general. Suppose, for example, that we wish to transmit the signal shown in figure 1.5 and we wish to preserve its general

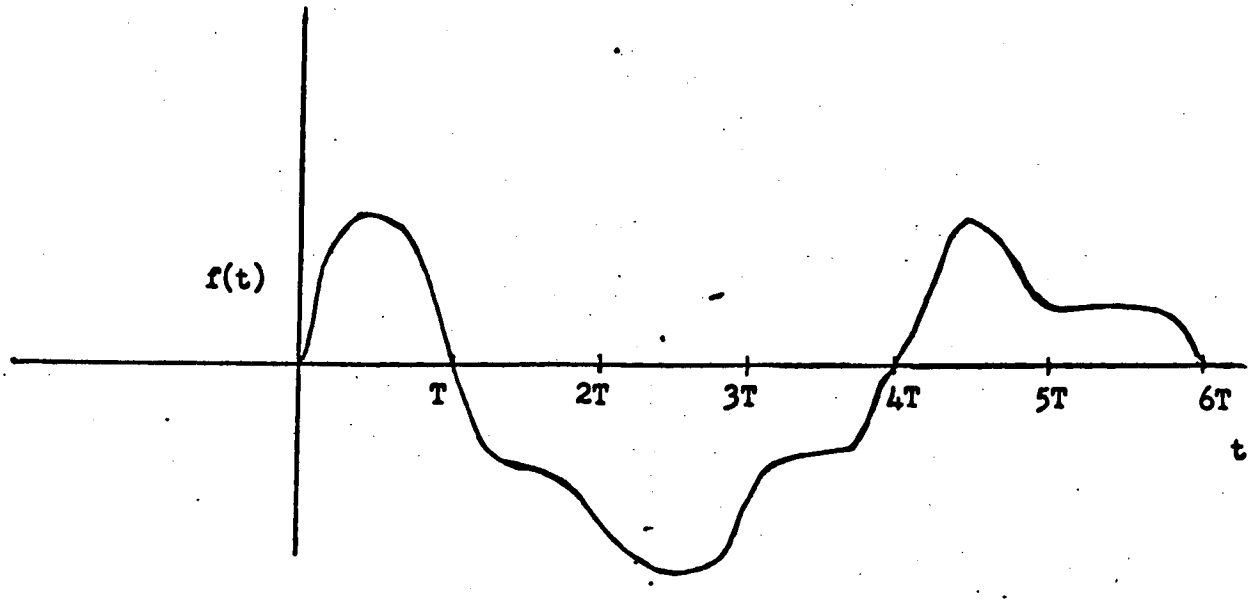


Figure 1.5. An Information Signal

shape. Although it may continue indefinitely let us consider only the interval $0 < t < 6T$, and, for convenience, let us assume that f(t) continues to generate this exact same waveform in the next 6T seconds, and so on for each succeeding 6T second period. Then f(t) would be periodic and we could expand it in a Fourier series, and by truncating this series at some point, we can estimate the bandwidth that will be necessary to transmit f(t). But observe that f(t) exhibits variations similar to a sinusoid with the period 2T. It should be expected (and, indeed, could be verified) that unless at least the frequency components out to the frequency $f_0 = \frac{1}{2T}$ were allowed to remain, the truncated series would be a poor approximation to the original function. This relationship holds in general. If it is desired to preserve in a time function variations in time of on the order of T seconds, it is necessary to allow it to occupy a bandwidth of at least $B = \frac{1}{2T}$ cps. Conversely, unless the bandwidth of a signal is at least B cps, variations of the signal occurring over a time duration of less than $\frac{1}{2B}$ seconds will not be observed.

1.1.6 Signal-to-Noise Ratio

An important parameter in the evaluation of any communication system is the signal-to-noise ratio of the received signal. This is defined as the ratio of the power in the information portion of the received signal to that in the portion due to the system noise. Specifically, suppose the received signal y(t) can be expressed as

$$y(t) = s(t) + n(t) \qquad (1.10)$$

where s(t) is due to the transmitted signal, while n(t) is that part which would exist even if no signal had been transmitted. Then the signal-to-noise ratio is defined as

$$\frac{S}{N} = \frac{\lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} s^2(t)dt}{\lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} n^2(t)dt} \qquad (1.11)$$

$$= \frac{\lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} s^2(t)dt}{\lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} (y^2(t)-s(t))^2dt}$$

(Note that we must now take the limit as T becomes infinite to determine the average power in a non-periodic function). This second expression for the signal-to-noise ratio can be regarded intuitively in the following sense: The power in the information portion of the signal can be thought of as a measure of the uncertainty of this information. The wider the range over which it varies, the greater is its power, and the less certain is the receiver of this information as to the value it will assume at some future time. The numerator represents, in the same sense, the uncertainty in the signal after it has been received, for the difference between the received signal y(t) and the information s(t), is a measure of the lack of certainty on the part of the user concerning this information. Then, the signal-to-noise ratio can be interpreted as the factor by which the uncertainty in the information is reduced by the

reception of the information bearing signal. This interpretation of the signal-to-noise ratio is one worth bearing in mind in the subsequent investigations.

As we have already observed, the power in a signal can also be evaluated in terms of its power spectral density when its bandwidth is known. The power spectrum of the noise with which we shall be concerned is constant at all frequencies (i.e., is white). Thus, if such noise, with a constant power spectral density N_0 watts/cps were applied to a filter which passed all frequencies in the range $f_0 \leq f \leq f_0 + B$ without any attennation and blocked all frequencies outside this range, the noise power at the output of this filter would be N_0B watts. Although filters with this property (often called ideal filters) cannot be constructed, it is often convenient to treat the receiver as though it were such a filter so far as the noise is concerned. Thus, the noise bandwidth B_n of a receiver is defined as the bandwidth of an ideal filter necessary to account for the noise power actually observed at its output. The noise at the output of a receiver is completely specified, therefore, in terms of its noise bandwidth and its effective noise temperature.

2 AMPLIFIERS AND ANTENNAS

As we observed in the previous chapter, noise is an unavoidable part of any communication system. In space telemetry systems this noise is essentially additive, white, and Gaussianly distributed. Since, as we argued earlier, it is the ratio of the signal to the noise power that determines the performance of any communication system , the same result can be accomplished by either increasing the signal power at the receiver, or decreasing the noise power. In this chapter we discuss ways for accomplishing both of these tasks.

The greatest noise contribution, in most space telemetry situations arises in the initial stages of the receiver. Since the transmitter operates at relatively high signal levels, the signal-to-noise ratio at the transmitter can be kept very large. Background radiation at the frequencies generally used is relatively insignificant (and, in any event, unavoidable.) At the receiver, however, the signal power is extremely low so that any noise contributed in the initial process of amplifying this signal may be, in comparison, most significant. It is at this stage that the greatest effort is demanded to decrease the additive noise, and it is here that the most spectacular progress has been made.

2.1 Low Noise Amplifiers

Signal amplification is most commonly achieved with vacuum tube and transistor amplifiers but because such amplifiers are relatively noisy and, at any rate, not practical at the frequency ranges used for

space communication, we shall not be concerned with them here. The earliest technique for the low-noise amplification of microwave frequencies involved the use of the traveling-wave tube. The traveling-wave tube, developed during World War II, relies upon the interaction between an electron beam and the signal-bearing electromagnetic wave. This electromagnetic wave is effectively slowed down to the velocity of the electron beam by passing it through a wave guide, generally in the shape of a helix. Since electromagnetic energy traverses linearly along a waveguide at nearly the velocity of light c, its rate of progress along the axis of the helix is approximately $(\ell/L)c$, where ℓ is the length of the axis of the helix, and L is the length of the wave guide comprising the helix. Thus, it is possible to make the velocities of linear propagation of the signal and the electron beam equal. When this is done, there is an interaction between the electric field of the signal and the electrons in the beam. The electrons densities are increased or decreased depending upon the intensity and direction of the field. This "bunching" in turn causes the field to be intensified in proportion to its original strength, thus producing amplification. Extremely large amplifications over a wide band of microwave frequencies are, indeed, possible with this technique. The noise arises, as usual, because the electrons do not all have the same energy or velocity. Thus the bunching cannot be perfect. Since the electrons are not all moving with the same velocity, they exhibit a counter-tendancy toward a random distribution. This appears as noise at the output. Much effort has been made to decrease the noise inherent in traveling wave tubes, and amplifiers

using these tubes have been built with effective noise temperatures of less than 300°K.

Another more recent development in low-noise broad-band microwave amplifiers is the parametric amplifier. The action of this device is commonly compared to the method by which a child, sitting in a swing, is able to increase the amplitude of this swinging arc. At the height of his displacement, when the swing changes directions, the child pulls back on the ropes, thereby slightly increasing his height, and hence his potential energy. At the bottom of the arc, the tension on the ropes is relaxed so that this potential energy is entirely converted into kinetic energy. Because the maximum height was increased this kinetic energy is greater than it would have been and, at the next peak, the potential energy has increased over its value at the previous peak. The energy of the child, therefore, is converted into oscillation energy of the swing.

In the same way any oscillator or resonant device can gain energy by being "pumped" at the right times. In fact, it can be shown that the oscillator exhibits a net energy gain even if it is pumped at the wrong time, too, so long as at least some of the pumping occurs at the peaks and troughs of the potential and kinetic energy storage cycles.

This, then, is the principle of the parametric amplifier. An oscillator or resonator possesses, generally, two means of energy storage. If the storage capacity of one of these devices (or parameters) is altered (as the child changes the effective length of the ropes of the swing), at a frequency high compared to its natural frequency, the resonator will exhibit a net increase of energy. This increase can be sizeable.

One realization of the parametric amplification principle is obtained through the use of solid state devices which have the property that their energy storage capacity varies in inverse proportion to the intensity of the applied signal. This applied, or pumping, signal is chosen to have perhaps ten times as great a frequency as the signal to be amplified. The latter signal, then, when used as in input to the resonant circuit containing the pumped element, is thereby amplified. Noise in such a device arises primarily from the disaptive elements in the storage devices and the associated circuitry. These effects can be kept to a minimum, however, by cooling the amplifier, with liquid helium, for example, to a few degrees Kelvin. It is possible, in this way, to get effective noise temperatures of 100°K or less for parametric amplifiers.

Another implementation of the parametric principle is the electron beam parametric amplifier. Here, an electron beam is used as the energy storage device and is pumped by an electric field. The performance of this device is approximately that of the more conventional parametric amplifier.

Probably the most successful low-noise amplifier yet developed, however, is the maser (acronym for microwave amplification by stimulated emission of radiation). The electrons in the crystal lattice of any material, like all electrons, spin about some axis. The orientation of the spin axes are restricted to certain positions, and normally the vast majority of the electrons are in the lowest energy position. If the difference in energy between the lowest two energy levels is ΔE, an amount of energy ΔE is absorbed by the crystal when an electron

makes a transition from the lowest to the next lowest level, and an amount of energy ΔE is radiated when the reverse transition occurs. Normally transitions occur equally often in each direction so that the net radiated energy is zero. The frequency of this radiation, we know from Planck's equation, must be

$$f = \frac{\Delta E}{h} \tag{2.1}$$

where $h = 6.6 \times 10^{-34}$ joule-seconds. If the crystal is radiated with energy at the frequency $\frac{\Delta E}{h}$, electrons are caused to make the transition to the higher level and energy is absorbed. By irradiation with energy at a higher frequency f' it is possible to excite the electrons to a still higher energy level $\Delta E' = hf'$. By the proper selection of a crystal it is possible to achieve a situation in which electrons, excited to the level $\Delta E'$ can decay to the level ΔE, but cannot decay further, to the ground level except in the presence of external radiation at the frequency $f = \frac{\Delta E}{h}$. It is thus possible to create a situation in which the majority of electrons are at the next to the lowest energy level. When this is the case, a signal at the frequency f when applied to the crystal exhibits a net increase in energy due to the preponderance of electron transitions to the ground level which it triggers. Thus, energy is transferred from the higher excitation frequency to the crystal, and from the crystal to the lower signal frequency. Again, the resulting amplification can be sizeable.

The noise generated in a maser amplifier can be exceedingly small. It is due to fluctuations in the radiation field in the neighborhood of the crystal. These fluctuations can be caused by thermal agitation

of the electrons causing a noise spectral density $N_{th} = kT$ where T is the actual temperature (in degrees Kelvin) of the crystal. In addition, however, radiation is emitted due to spontaneous electron transitions which, according to quantum mechanics, give rise to noise with a spectral density $N_q = (1/2)hf$. For microwave frequencies the total noise $N_0 = N_{th} + N_q$ can be quite small and masers have been built with an effective noise temperature of less than 10°K. Note, however, that the N_q term is independent of temperature and hence cannot be reduced by cooling the crystal. This term, being proportional to frequency, becomes more significant at higher frequencies. At frequencies in the visible light range, for example (masers which operate at frequencies approximating those of visible light are called lasers, the m of microwave becoming the l of light) the effective noise temperature can increase to as much as 20,000° to 30,000°K, thus seriously counteracting some of the real advantages associated with the use of lasers in space communications.

2.2 Antenna Gain

Another method for mitigating the severe conditions encountered in space communications due to the vastly increased distances between the transmitter and receiver is through the use of high gain antennas. The gain of an antenna is defined as the ratio of the maximum power intensity Φ_m (the amount of power incident upon a unit area) in any direction to the average intensity Φ_{ave}. Usually, the gain G is expressed in decibels:

$$G = 10 \log_{10} \left[\frac{\Phi_m}{\Phi_{ave}} \right] \text{ db.} \qquad (2.2)$$

Since we are interested in communication from a point to a point rather than from a point to many points, as in commercial radio, we clearly

want a somewhat different antenna design than that commonly used in the latter case. In conventional radio transmission it is desired to radiate equally in all horizontal directions. To accomplish this verticle or "dipole" antennas are used with heights which are, ideally, half of the wave length of the frequency radiated. Since they radiate horizontally, with little energy being transmitted vertically, they exhibit gains which are greater than 1; in the case of an ideal dipole antenna the gain in the equatorial plane is 2.15 db.

For space communication, however, it is desired to radiate energy only between the one transmitter and the one receiver. (There may be more than one receiver in practice but, at deep space distances, the earth itself is effectively one point.) It is therefore necessary to be able to direct the radiated energy in a narrow beam toward the receiver. This is most effectively accomplished by focusing the energy by means of a reflector in the shape of a parabaloid. A parabola, it will be recalled, has the geometrical property illustrated in figure 2.1 and hence, to the extent that the angle of incidence of a microwave beam is equal to the angle of reflection, all energy originating at the focal point and striking the antenna, will be reflected in a direction parallel to the axis of the antenna A-A'.

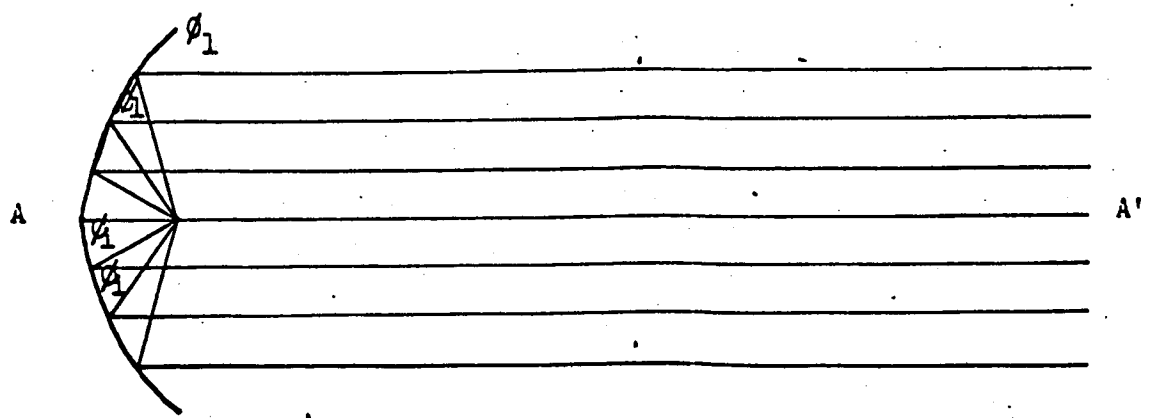


Figure 2.1 Direction of Reflection from a Parabolic Antenna

Unfortunately, however, the wavefront does not remain constant with a diameter equal to that of the antenna; rather, it increases in area. For an intuitive understanding of the reason for this spread consider the illustration in figure 2.2.

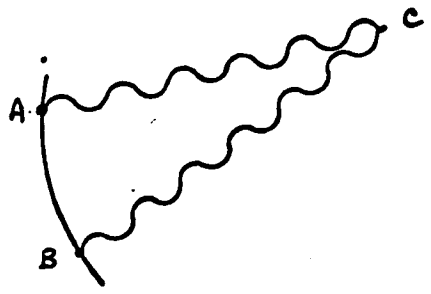


Figure 2.2. Interference Due to Reflections from Different Parts of the Antenna

The observer at point C "sees" energy reflected from various portions of the surface of the antenna. Consider two infinite small signal surface areas A and B. Since C is closer to A than to B, the energy from B must travel farther before it reaches the point C. If the geometry is such that the distance BC is exactly one-half wave-length further than the distance AC, then the radiation from the two points A and B will arrive at C exactly 180° out-of-phase. The electromagnetic fields will have equal amplitudes but opposite signs and will, therefore, completely cancel each other. When C is too close to the axis, there will be no two points on the surface of the antenna such that the difference in their distance to C is as great as one-half wave-length. As the distance from C to the axis

increases, the points A and B satisfying the property describe will move closer together, and in addition, other points A' can be found on the surface of the antenna such that the di A'C and B'C differ by exactly three-halfs wave-length. Thus, as C moves away from the axis there is more and more cancellation, so that the net amount of energy striking C rapidly diminishes. To estimate the width of the beam at a distance D from the antenna consider the diagram in figure 2.3.

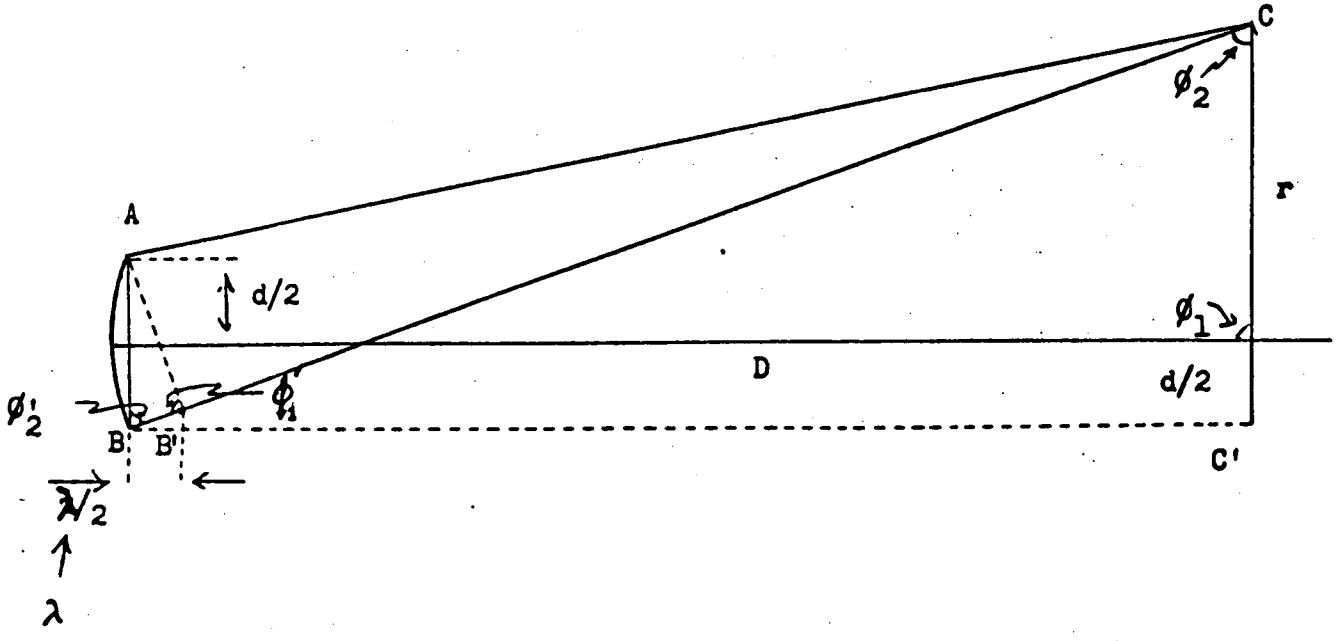


Fig. 2.3 Determination of the Beam Width

We want to find as a measure of the beam width, the smallest value of r, the distance from C to the axis, such that there is total cancellation of energy arriving from at least two points on the antenna. Clearly, the first two points on the antenna surface which provide such cancellation are those two points in the same plane as C, and separated by the maximum distance. Therefore, let A be at one extreme of the antenna and B at the other, separated from A by a distance d, the

diameter of the antenna. Assuming that D and r are large compared to the dimensions of the antenna, and that λ is small compared to d, it is easy to determine the value of r in terms of the wavelength, the antenna diameter, and the distance D. First we find the point B' on the line CB such that the distances CA and CB' are equal. Since D is large, ϕ_1' is nearly a right angle and hence $\phi_1' \doteq \phi_1$. Clearly, $\phi_2' = \phi_2$ and, hence, the triangles BC'C and AB'B are (nearly) similar. Thus

$$\frac{AB'}{BB'} = \frac{BC'}{CC'}$$

and since $AB' = (d^2 + \frac{\lambda^2}{4})^{\frac{1}{2}} \simeq d$, $BB' = \frac{\lambda}{2}$, (in order that we get the desired cancellation), $BC' = D$, and $CC' = r + d/2 \simeq r$, we have

$$r \simeq \frac{D\lambda}{2d} \tag{2.3}$$

and the beam width is proportional to $\frac{D\lambda}{d}$.

Now consider the amount of power received by a second parabolic antenna of area A_R at a distance D from the first. Since the beam width is proportional to 2r, its area is proportional to $(2r)^2$ and hence to $\frac{D^2\lambda^2}{d^2}$. The percentage of the power which is received, assuming $A_T < \frac{D^2\lambda^2}{d^2}$, is clearly proportional to the ratio of the area of the receiving surface to the area of the beam, since all the power striking the antenna surface is reflected to the focal point (the antenna is assumed to be parabolic) and hence to the receiver input. Consequently, designating by P_T the total transmitted power, and by P_R the total received power, we have

$$P_R \propto P_T \cdot \frac{A_R}{\frac{D^2\lambda^2}{d^2}}$$

And finally, since the area of the transmitting antenna is proportional to d^2 we have

$$P_R \propto P_T \frac{A_R A_T}{\lambda^2 D^2}$$

Actually, this heuristically derived result can be shown to be exact, so that

$$P_R = P_T \frac{A_R A_T}{\lambda^2 D^2} \tag{2.4}$$

For non-ideal parabolic antennas, A_R and A_T must be replaced by an effective area which is always somewhat less than the true area. This is primarily because of the fact that it is difficult to radiate the entire surface of the antenna with energy of equal magnitude and equal phase. Typically, the affective area is 50% to 80% of the actual area.

As shown, the beam area at a distance D from the antenna is proportional to $\frac{D^2\lambda^2}{A_T}$. The power intensity over the beam front is therefore proportional to $\frac{P_T A_T}{D^2 \lambda^2}$. If the power were radiated uniformly in all directions, the power intensity at a distance D from the antenna would be equal to $\frac{P_T}{4\pi D^2}$ since the power is uniformly distributed over a spherical surface of area $4\pi D^2$. Recalling the definition of gain, we have

$$G_T = 10 \log_{10} \left(\frac{\phi_m}{\phi_{ave}}\right) \propto 10 \log_{10} \left[\frac{4\pi A_T}{\lambda^2}\right] \tag{2.5}$$

Again, this equation can be shown to be exact so long as A_T is interpreted as the effective area. Similarly, the gain of the receiving parabolic antenna is

$$G_R = 10 \log_{10} \left[\frac{4\pi A_R}{\lambda^2}\right] \tag{2.6}$$

Then the received power in decibels is

$$10 \log_{10} P_R = 10 \log_{10} P_T + G_R + G_T - 20 \log_{10} \frac{4\pi D}{\lambda} \qquad (2.7)$$

Equation (7) is valid regardless of the type of antennas used so long as G_R and G_T are the appropriate antenna gains.

In order to maximize the amount of power received, or equivalently, to maximize the gains of the two antennas, it is necessary to make the parabolic antennas as large as possible and the wave lengths as short as possible. First of all, there are practical limitations to the shortness of the wave length. One of these limitations stems from the fact that, beyond a certain point, the effective noise temperature of the best amplifiers increases sharply as the wavelength decreases, thereby counteracting the advantages in antenna gain. But more critically, in order to realize the theoretical gains of parabolic antennas, the dimensions of the antenna must be accurate to within a fraction of the wavelength. Since the gain increases in proportion to the area, it is advantageous to make the area as large as possible. But the larger the area, the more difficult it is to keep the tolerances within the necessary limitations. Thus, there is a trade-off between the area and the wavelength. Moreover, because the transmitting and receiving antennas are moving with respect to each other, it must be possible to move the ground based antenna so that this also places restrictions on its size, (the fact that the vehicle antenna must be propelled through space, of course, limits its size). Finally, the transmitter antenna must be pointed in space with an accuracy proportional to the width of the beam

or the maximum energy is not received at the receiver antenna. This clearly also limits the gain, and becomes particularly significant at very short wavelengths.

Other than parabolic antenna designs are also sometimes used in space telemetry. An omni-directional antenna which, ideally, radiates or receives energy equally in all directions is always included on a spacecraft as a safety factor to enable transmission to and from the vehicle regardless of its orientation in space. Evidently this antenna has unity gain (0 db) in all directions.

Clearly, stationary antennas can be made much larger than those that must be moved. Because space telemetry antennas must be accurately pointed in space, stationary antennas are not too useful for this purpose. It is possible to get some effective direction change in stationary antennas by properly controlling the position of the source which radiates the antenna as well as the relative phases of the energy striking the various parts of the antenna surface. Such antennas are particularly useful in radio astronomy.

Other stationary, or partially moveable antennas, have been constructed which rely for their gain upon the spacing of the various component parts. The Mills cross is an example of such a technique. Here, two long horizontal antennas are arranged in a cross. The resolution of a Mills cross (i.e., its ability to receive energy from only a very narrow region of space, a factor especially significant in radio astronomy) is comparable to that of an antenna whose area is equal to the product of the length of the two legs of the cross. Other crosses and arrays consisting of a number of dipoles or parabolic antennas suitably placed have been found to have large gain properties with mechanical tolerances which are still within reason.

3 ANALOGUE MODULATION

Before concentrating on some of the recent modulation techniques for space communications it is well to review the more conventional methods of wireless long-distance communications. Typically a signal of the form $\sqrt{2}\ B \sin(\omega t + \phi)$ is generated at the transmitter. If the frequency $f = \frac{\omega}{2\pi}$ is sufficiently high, this signal can be applied to an antenna and will cause an electromagnetic wave to be emitted into space. A signal $\sqrt{2}\ A \sin\left[\omega(t-\tau) + \phi\right]$ will then be excited at the receiver antenna, where $k = \frac{A}{B}$ represents the attenuation due to the medium and the distance through which the signal travels, and τ the delay representing the time needed for the signal to travel from the transmitter to the receiver. (The medium here and throughout these notes is assumed to be constant and to involve only one transmission path. In particular, k does not vary with time, except perhaps for a slow steady change due to a change in the distance between the transmitter and the receiver.)

If B and ω and ϕ are kept constant at the transmitter, virtually no information can be transmitted. The receiver is able to determine that the transmitter must exist, but essentially nothing else. If, on the other hand, any one, or a combination of these parameters is varied in accordance with some rule known at both the transmitter and receiver, information can be transmitted. Commonly there is some time function f(t) which is to be transmitted representing, for example,

a temperature or pressure reading on a space vehicle, or a sound or light intensity in commercial broadcasting. Thus if A in particular is made to vary proportionately with f(t), A(t) = af(t), the resulting amplitude modulated signal is capable of conveying information. Similarly, if $\phi(t) = bf(t)$ or if $\omega(t) = cf(t)$ the signal is said to be phase modulated and frequency modulated respectively. These three types of modulation will be examined in some detail in the next few sections.

3.1 Amplitude Modulation

The generation of an amplitude modulated signal is relatively straightforward. The signal f(t) is converted to a voltage intensity in accordance with its amplitude (e.g. a sound wave is passed through a microphone). This voltage is amplified and multiplied by a signal $\sqrt{2}\sin(\omega_0 t + \phi)$. In addition, for reasons which will become apparent shortly, some unmodulated signal is also added in. All of these procedures can be readily accomplished electronically. The resulting amplitude modulated signal $x(t) = \sqrt{2}\,B(1+mf(t))\sin(\omega_0 t+\phi)$ is then transmitted. The parameter m is referred to as the modulation index.

The spectrum of the signal x(t) is simply a frequency translation of the spectrum of the modulating signal f(t). To illustrate this, let us suppose that f(t) is a sinusoidal signal:

$$f(t) = a \sin \omega_n t. \qquad (3.1)$$

Then

$$x(t) = \sqrt{2}\, B\, (1 + am \sin \omega_n t) \sin (\omega_0 t + \phi)$$

$$= \sqrt{2}\, B \sin (\omega_0 t + \phi) + \frac{\sqrt{2}}{2}\, B\, am \cos ((\omega_0 - \omega_n)t + \phi) \qquad (3.2)$$
$$- \frac{\sqrt{2}}{2}\, B\, am \cos ((\omega_0 + \omega_n)t + \phi)$$

and a modulating signal at the frequency $f_n = \frac{\omega_n}{2\pi}$ results in a modulated signal having power at the carrier frequency $f_0 = \frac{\omega_0}{2\pi}$, at the sum frequency $f_0 + f_n$ and at the difference frequency $f_0 - f_n$.

The power spectrum of the modulated signal is shown in figure 3.1. If f_M were the maximum frequency and f_m the minimum frequency of the modulating signal (as shown in figure 3.2(a)) the frequency range of the

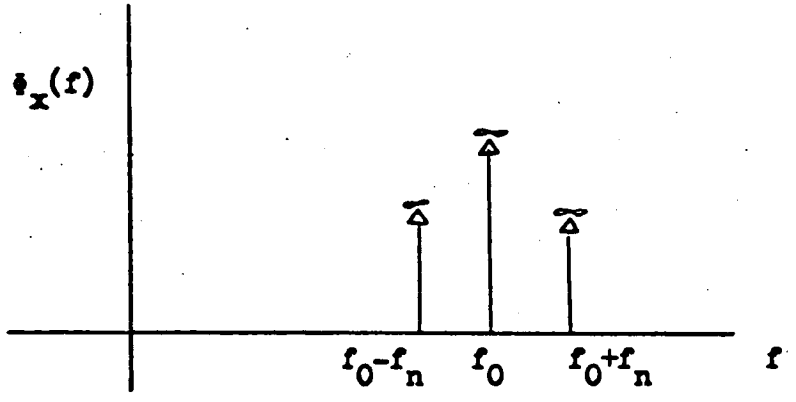


Figure 3.1. Power Spectrum of a Carrier of Frequency f_0 Amplitude Modulated by a Sinusoid of Frequency f_n

modulated signal would extend from $f_0 - f_M$ to $f_0 - f_m$ and from $f_0 + f_m$ to $f_c + f_M$ (as shown in figure 3.2(b)). The bandwidth of the modulated signal is consequently double that of the modulating signal.

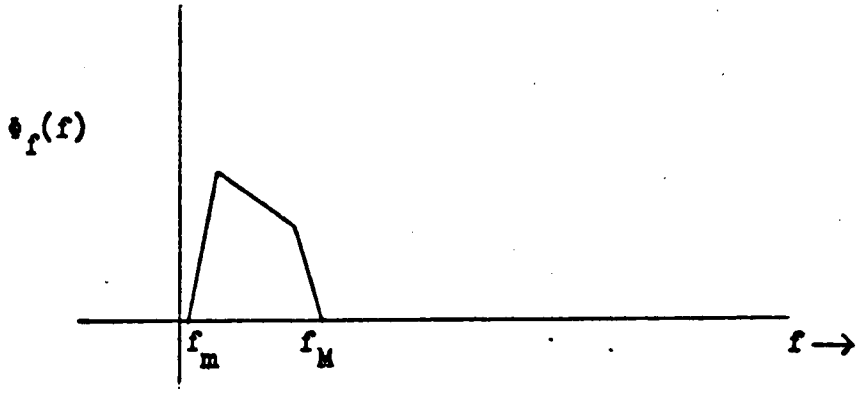
Again, if the modulating signal were a sinusoid, the average transmitted power would be

$$P_T = \lim_{T\to\infty} \frac{1}{T} \int_{\frac{-T}{2}}^{\frac{T}{2}} x^2(t)dt = \frac{2B^2}{\frac{2\pi}{\omega_0}} \int_0^{\frac{2\pi}{\omega_0}} \sin^2(\omega_0 t+\phi)tdt +$$

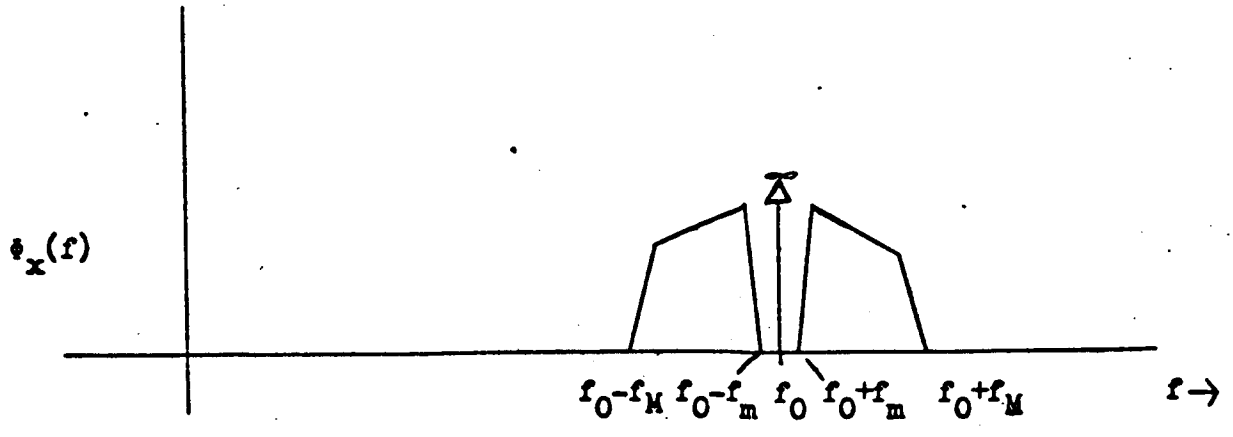
$$+ \frac{B^2 a^2 m^2}{2(\frac{2\pi}{\omega_0-\omega})} \int_0^{\frac{2\pi}{\omega_0-\omega_n}} \cos^2((\omega_0-\omega_n)t+\phi)dt +$$

$$+ \frac{B^2 a^2 m^2}{2\frac{2\pi}{\omega_0+\omega_n}} \int_0^{\frac{2\pi}{\omega_0+\omega_n}} \cos^2((\omega_0+\omega_n)t+\phi)dt$$

$$= B^2 + \frac{B^2 a^2 m^2}{2} \qquad (3.3)$$



(a) Modulating Signal Power Spectrum



(b) Modulated Signal Spectrum

Figure 3.2. Spectrum Translation Due to Amplitude Modulation

where $x(t)$ is as defined in equation (2). Note that the cross-product terms all vanish since they involve the products of sinusoids at different frequencies. The first term in the equation for P_T represents the carrier power while the second corresponds to B^2m^2 times the power in the modulating signal. For a general modulating signal $f(t)$ the expression for the average transmitted power is

$$P_T = B^2 + B^2 m^2 P_m \tag{3.4}$$

where P_m is the average power in the modulating signal. The percentage of the total power in the modulation is

$$\% \text{ of power in modulation} = \frac{B^2m^2P_f}{B^2+B^2m^2P_f} = \frac{m^2P_f}{1+m^2P_f} \tag{3.5}$$

3.2 Demodulation of AM

In order to obtain useful information from the signal $x(t)$ at the receiver it is necessary to <u>demodulate</u> it to obtain the desired signal $f(t)$. An AM signal may be demodulated in a number of ways. The most common technique involves the use of a non-linear element called a half-wave rectifier followed by a filter. The ideal half-wave rectifier can be regarded as a device whose output $y(t)$ is related to the input $x(t)$ as follows:

$$y(t) = \begin{cases} x(t) & x(t) \geq 0 \\ 0 & x(t) < 0 \end{cases} \tag{3.6}$$

A typical amplitude modulated waveform is shown in figure 3.3. It is seen that so long as $mf(t) > -1$, the signal $x(t)$ is positive for $2kT < t < (2k+1)T$ and negative for $(2k+1)T < t < (2k+2)T$, for all integer values of k and for $T = \frac{2\pi}{\omega_o}$.

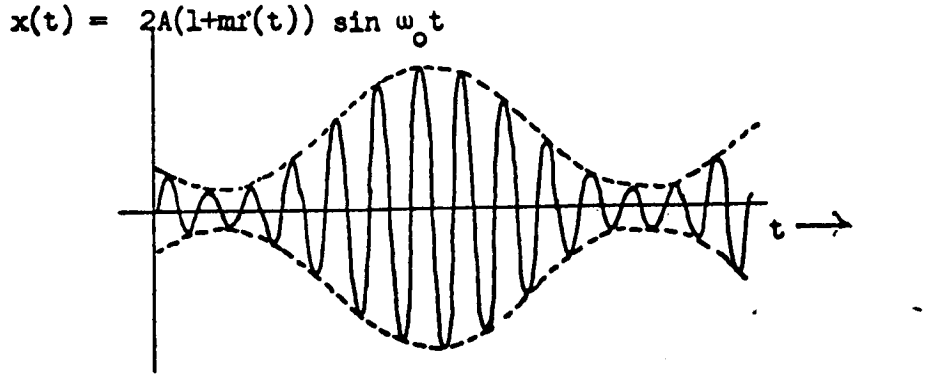


Figure 3.3 An AM Signal

Consequently, the half-wave rectifier has the same effect on the received waveform as if it were multiplied by the square wave shown in figure 3.4.

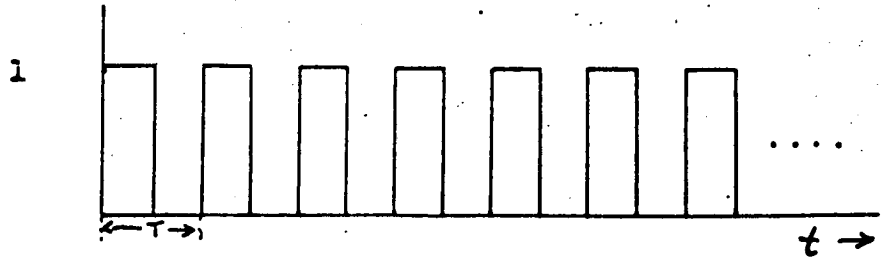


Figure 3.4 A Square wave of period $T = \frac{2\pi}{\omega_o}$

But, as shown in section 1.1, the Fourier series expansion of this square wave s(t) is

$$s(t) = \frac{1}{2} + \sum_{n=1}^{\infty} \frac{\sin \frac{\pi n}{2}}{\frac{\pi n}{2}} \sin n\omega_o t \qquad (3.7)$$

Then

$$x(t)\, s(t) = \sqrt{2}\, A(1 + mf(t)) \sin \omega_o t \left[\frac{1}{2} + \sum_{n=1}^{\infty} \frac{\sin \frac{\pi n}{2}}{\frac{\pi n}{2}} \sin n\omega_o t \right]$$

$$= \sqrt{2}A(1+mf(t)) \left\{ \frac{1}{2} \sin \omega_o t + \sum_{n=1}^{\infty} \frac{\sin \frac{\pi n}{2}}{\frac{\pi n}{2}} \frac{1}{2} \left[\cos(n-1)\omega_o t - \cos(n+1)\omega_o t \right] \right.$$

which is just a sum of amplitude modulated signals. (3.8)

The product x(t) s(t), therefore, contains terms centered about the frequencies 0, f_o, $2f_o$ $4f_o$, ... as shown in figure 3.5.

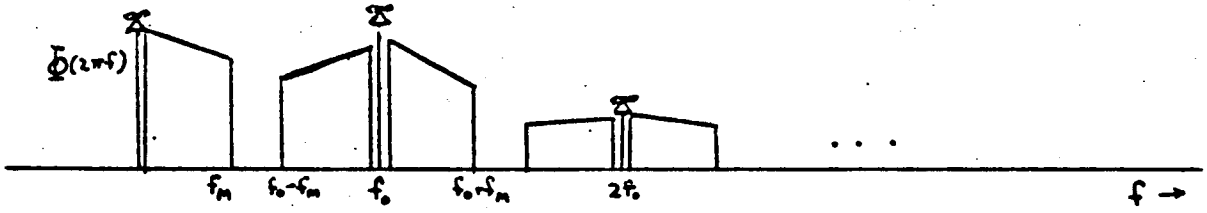


Figure 3.5 The Spectral Density of the Product x(t) s(t).

It is seen that all but the desired term $\frac{\sqrt{2}\,A}{\pi}(1+mf(t))$ can be eliminated by filtering if $f_o - f_M > f_M$ or if $f_M < f_o/2$ where f_M is the highest frequency in the modulating signal.

3.3 Double- and Single-Sideband Modulation

While the method of AM communications described in the previous section is quite satisfactory for commercial use it has some important limitations in those situations in which the available power is limited. First it will be recalled, the demodulation scheme outlined requires that $mf(t) > -1$. To appreciate the signifi-

cance of this limitation, suppose $f(t) = \sin 2\pi f_m t$. Then $P_f = \frac{1}{2}$, and since $mf(t) > -1$, m must be less than one, and the percentage of power in the modulation is less than 33-1/3% (cf. equation 5).

To overcome this difficulty, consider the following demodulation scheme: A narrow-band filter is centered about f_o and the output is used to estimate the frequency and phase of the carrier.

Suppose that the carrier is of the form $\sin \omega_o t$ and the estimate $c(t) = \sqrt{2} \sin(\omega_o t + \theta)$ is made where presumably θ is small. Then the signal can be demodulated by forming the product

$$x(t)\, c\,(t) = 2A\left[1 + mf(t)\right] \sin \omega_o t \sin(\omega_o t + \theta)$$

$$= A\left[1 + mf(t)\right]\left[\cos \theta - \cos(2\omega_o t + \theta)\right]$$

which, after filtering, yields the desired signal

$$A\left[1 + mf(t)\right] \cos \theta \tag{3.9}$$

Note that no limitations have been placed on the maximum value of $mf(t)$. The only requirement now is that there is enough power in the carrier to enable a good estimate of its frequency and phase. It is not obvious that this is superior to the previous AM system until we determine how much power must be included in the carrier for satisfactory results. In section 3.8 we shall verify however, that in a typical situation, less than 1% of the total power need be included in the carrier, thus allowing a substantial increase in performance over conventional AM. This technique of increasing the proportion of power in the modulation by suppressing the carrier is commonly referred to as double-sideband, suppressed-carrier modulation, DSB/SC.

An interesting modification of the double-sideband amplitude modulation system is afforded by a technique known as single-sideband modulation, SSB. Observe from the discussion in section 3.1 that the power spectrum $\Phi(\omega)$ of an amplitude modulated signal is symmetric about the carrier frequency. When f(t) is a sinusoid of frequency f_n the amplitude modulated signal contains a term at the frequency $f_0 + f_n$ and a term of the same amplitude at the frequency $f_0 - f_n$. This same symmetry occurs regardless of the signal f(t).

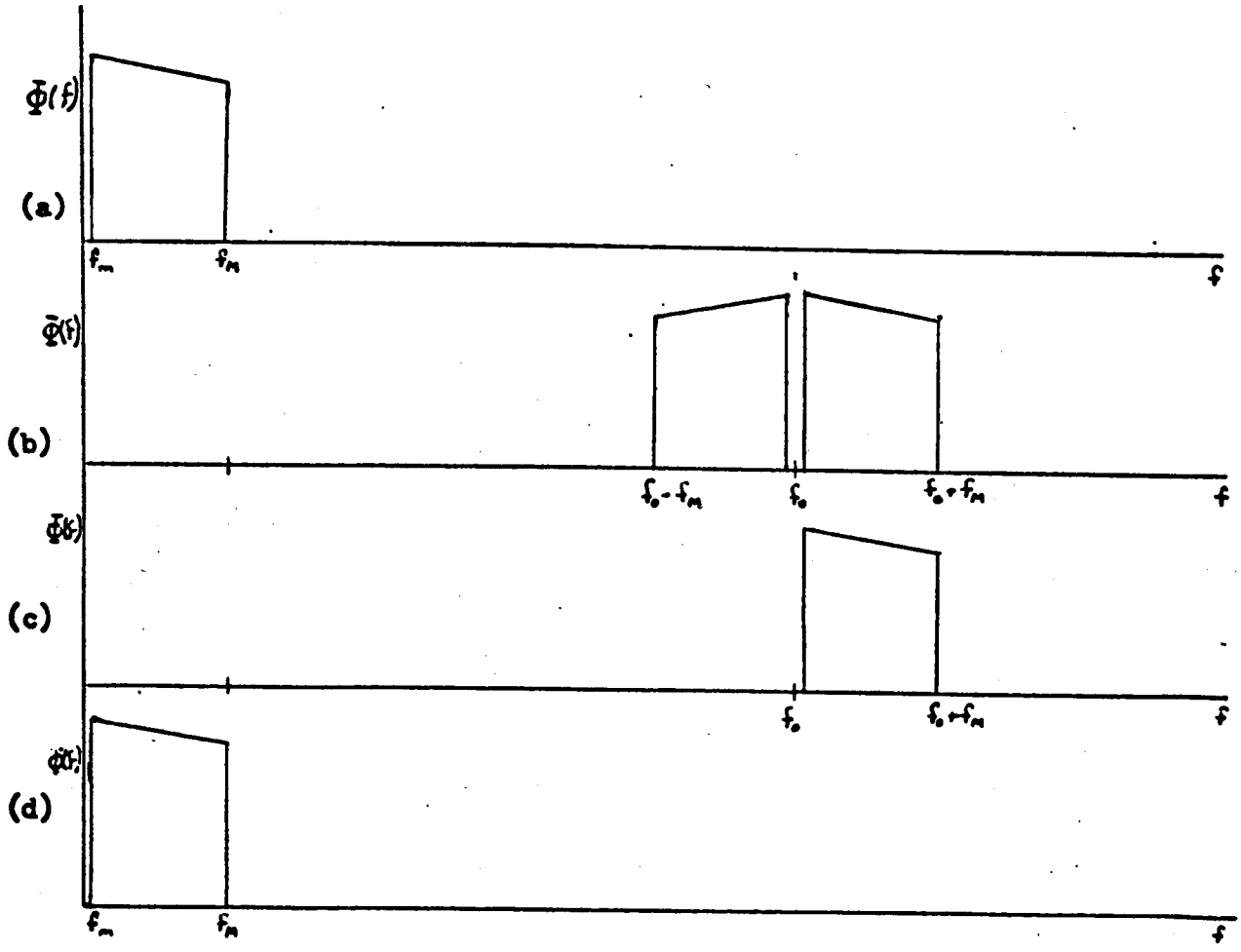


Figure 3.6 Power Spectra Illustrating the Philosophy of SSB

Conventional AM or DSB modulation involves the product $f(t) \sin \omega_o t$, which, as seen in section 3.1, simply shifts the spectrum of the modulating signal 3.6(a) to the position shown in figure 3.6(b). Now suppose the signal $f(t) \sin \omega_o t$ is passed through an ideal band pass filter with the pass band $f_o + f_\ell < f < f_o + f_m$. The output then has the spectrum illustrated in figure 3.6(c). If this filtered signal $f'(t)$ is transmitted and demodulated by forming the product $f'(t) \sin \omega_o t$ and filtering out the high-frequency components, the resulting signal has the spectrum shown in figure 3.6(d).

This signal is identical to the original modulating signal $f(t)$. But note that by filtering before transmission, as described, only half as much bandwidth is needed for SSB as for DSB modulation. As with DSB modulation, it is necessary to transmit some carrier power in order to demodulate a SSB signal. It can be shown that in the latter case, however, the phase accuracy of the estimate need not be as great as before to ensure the same performance.

3.4 Noise Analysis of Amplitude Modulation Communication

The ultimate evaluation of any communication system rests in its performance in the presence of noise. A convenient measure of this performance is the output signal-power-to-noise-power ratio S/N. In the case of DSB and SSB modulation combined with product demodulation, this ratio is readily determined. Consider first DSB* modulation. This received signal may be written

$$x(t) = \sqrt{2}\, a \sin \omega_o t + \sqrt{2}\, Af(t) \sin \omega_o t$$

The total power in the modulation is $A^2 P_f$ where P_f represents the power in the modulating signal. The signal $x(t)$ is demodulated by forming the product

* When we refer to DSB modulation here we intend double-sideband suppressed-carrier modulation. The DSB/SC designation is somewhat redundant. Non-suppressed-carrier modulation we shall denote "conventional AM".

$\sqrt{2}\, x(t) \sin \omega_o t$ and passing it through a low-pass filter with the cut-off frequency $B = f_m$. The output due to the signal is therefore

$$\left\{ \sqrt{2}\ Af(t) \sin \omega_o t \quad \sqrt{2} \sin \omega_o t \right\}_{l.f.} = Af(t)$$

where the subscript l.f. designates the low frequency components only. The output signal power is consequently $A^2 P_f$, the power in the modulation. The output noise signal is

$$n_1(t) = \sqrt{2}\, n(t) \sin \omega_o t$$

which simply represents a frequency translation of the noise n(t). Since the input noise is assumed white it remains white after the product is formed, and since the average power in the sinusoid is unity, the power spectral density of $n_1(t)$ is the same as that of n(t), i.e., N_o. The output noise power is consequently $N_o f_m$ and the output signal-to-noise ratio for DSB modulation is

$$\left(\frac{S}{N}\right)_{DSB} = \frac{A^2 P_f}{N_o f_m} = \frac{P_T - a^2}{N_o f_m} \tag{3.10}$$

where P_T is the total received signal power, $P_T = A^2 P_f + a^2$. When SSB modulation is used, although half the signal spectrum is suppressed, the other half can represent twice the power as before, keeping the total radiated power the same. After forming the product $\sqrt{2}\, x_{SSB}(t) \sin \omega_o t$ and filtering, as before, it is evident from figure 3.6(d) that the situation is identical to that for DSB modulation. Hence

$$\left(\frac{S}{N}\right)_{SSB} = \frac{P_T - b^2}{N_o f_m} \tag{3.11}$$

where b is the amplitude of the received unmodulated carrier. Since, generally, a^2 and b^2 can both be small compared to the modulation power

$$\left(\frac{S}{N}\right)_{DSB} \doteq \frac{P_T}{N_o f_m} \doteq \left(\frac{S}{N}\right)_{SSB} \tag{3.12}$$

The demodulation scheme using a half-wave rectifier, of course, will not achieve the performance indicated here. But, because we are considering modulation from the viewpoint of space communications and not as applied to commercial radio and television, we are concerned primarily with how well a particular modulation scheme can be made to work, not how well it works using inexpensive, mass-produced receivers. This spares us the considerably greater difficulty of analyzing the signal-to-noise ratios resulting from the use of more common demodulators such as the half-wave rectifier.

3.5 Angle Modulation

In this section we will consider communication systems in which the signal

$$\sqrt{2}\ B \sin \theta(t) \tag{3.13}$$

is transmitted, the angle $\theta(t)$ varying in accordance with the modulating signal. If we define the instantaneous frequency as the rate of change of the phase angle $\theta(t)$, then $\omega(t) = \frac{d\theta(t)}{dt}$. Note that this corresponds to the intuitive notion of frequency when $\theta(t) = \omega t + \theta_0$. When ω varies with time however, the intuitive definition of frequency is somewhat less clear.

A phase modulation system is one in which the phase angle $\theta(t)$ is allowed to vary with the modulating signal $f(t)$:

$$\theta(t) = \omega_c t + \theta_0 + \Delta\theta f(t). \tag{3.14}$$

Frequency modulation, on the other hand implies that the instantaneous frequency is made to vary with $f(t)$:

$$\omega(t) = \omega_c + \Delta\omega f(t) \tag{3.15}$$

But since $\omega(t) = \frac{d\theta(t)}{dt}$

$$\theta(t) = \int \omega(t)\, dt = \omega_c t + \theta_o + \Delta\omega \int f(t)\, dt \qquad (3.16)$$

and FM is essentially PM with the exception that the modulating signal in the latter is the derivative of that in the former. For this reason, the two types of modulation can be analyzed simultaneously so long as this difference is borne in mind.

As with any modulation scheme, one of the first considerations when an FM (or PM) signal is to be transmitted is that of its bandwidth occupancy.

Consider the case in which the FM modulating signal is a sinusoid

$$f(t) = \cos \omega_n t$$

Then

$$\omega(t) = \omega_c + \Delta\omega \cos \omega_n t$$

and

$$\theta(t) = \omega_c t + \frac{\Delta\omega}{\omega_n} \sin \omega_n t + \theta_o$$

and hence

$$x(t) = \sqrt{2}\ B \sin \left[\omega_c t + \beta \sin \omega_n t + \theta_o\right] \qquad (3.17)$$

where

$$\beta = \frac{\Delta\omega}{\omega_n}$$

is the FM <u>modulation index</u>. By a trigonometric identity

$$x(t) = \sqrt{2}\ B \sin(\omega_c t + \theta_o) \cos \left[\beta \sin \omega_n t\right]$$

$$+ \sqrt{2}\ B \cos(\omega_c t + \theta_o) \sin \left[\beta \sin \omega_n t\right] \qquad (3.18)$$

Suppose, for the moment, that β is small, say less than $\frac{\pi}{18}$. Then $\beta \sin \omega_n t$ is always less than $\frac{\pi}{18}$ radians or 10°, and, to a good approximation

$$\left.\begin{aligned} \cos\left[\beta \sin \omega_n t\right] &\doteq 1 \\ \sin\left[\beta \sin \omega_n t\right] &\doteq \beta \sin \omega_n t \end{aligned}\right\} \quad \beta < \frac{\pi}{18} \tag{3.19}$$

The frequency modulated signal then is approximately

$$x(t) \cong \sqrt{2}\, B \sin(\omega_c t + \theta_0) + \sqrt{2}\, B\, \beta \sin \omega_n t \cos(\omega_c t + \theta_0) \tag{3.20}$$

which, except for a 90° phase shift in the unmodulated carrier, is exactly the same as if the amplitude modulated signal

$$x_{AM}(t) = \sqrt{2}B\,(1 + \beta \sin \omega_n t)\cos(\omega_c t + \theta_0) \tag{3.21}$$

were transmitted. In general, if the modulating signal is f(t), the frequency modulated signal can be approximated, for small modulation indices β by the signal

$$x(t) \cong \sqrt{2}\, B \sin(\omega t + \theta_0) + \sqrt{2}\, B\, \Delta\omega \left(\int f(t)dt\right) \cos(\omega_c t + \theta_0) \tag{3.22}$$

(Note that $\beta = \frac{\Delta\omega}{\omega_n} = \frac{\Delta f}{f_n}$ must be small for all frequencies of the modulating signal f(t) for this approximation to be valid. Specifically

if f(t) has a frequency component $a_n \sin \omega_n t$, then $a_n \Delta\omega/\omega_n$ must be small. When this situation holds the modulation is referred to as narrow-band FM. This is because the bandwidth required is the same as that needed for conventional AM when the modulating signal is f(t). In particular, if the highest frequency component is f(t) is f_m, then the maximum frequency component of its integral is also f_m and the narrowband FM bandwidth is just $2f_m$, as it would be with conventional AM. When is increased, however, it will be seen that the FM bandwidth can be considerably greater than that necessitated by AM.

Suppose now that $\beta = \Delta f/f_n$ is large. Then the frequency deviation Δf is large compared to the modulating frequency f_n. In this case, the frequency being transmitted varies from $f_c - \Delta f$ to $f_c + \Delta f$, and, importantly, it varies between these extremes at a rate which is slow compared to the distance over which it varies. The transmitted signal could be approximated by a signal at the frequency $f_c + \Delta f$, followed some time later by a signal $f_c + \Delta f - \delta f$, followed still later by $f_c + \Delta f - 2\delta f$, etc. The important feature of this signal is the frequency extremes through which it varies, not the relatively slow rate at which it changes frequencies. This suggests that the bandwidth required for FM, when β is large, is approximately $2\Delta f = 2\beta f_n$ instead of the value $2f_n$ needed when β is small. When β is on the order of ten or greater this estimate of the required bandwidth is quite accurate, and it is reasonably applicable even when β is as small as four. Since, when β is large, the FM bandwidth is increased by a factor of β over that needed for AM, this modulation scheme is designated wide-band FM.

Note that since $\beta = \Delta_\omega/\omega_n = \Delta f/f_n$ in the case of frequency modulation, $W = 2\Delta f$ when β is large and is independent of frequency so long as the amplitude Δf of the modulating signal does not vary with frequency. Since, as we

shall show, the performance of FM is proportional to its bandwidth, it is desirable to have maximum bandwidth occupancy as consistently as possible.

With a phase modulated signal, the analysis is identical except that now $\beta = \Delta\theta$ and the bandwidth is $f_n\Delta\theta$. Thus, if the amplitude of the modulating signal is independent of frequency, the bandwidth of a phase modulated signal increases with the modulating frequency, a generally less desirable situation. On the other hand, ordinary speech and music exhibit the property that the amplitude of a frequency component, beyond a certain frequency, tends to be inversely proportional to the frequency. In this case, $\Delta\theta \alpha\ 1/f_n$ and the bandwidth W of a PM signal remains constant, independent of frequency, whereas an FM bandwidth would decrease with increasing frequency. For this reason, commercial FM modulating signals are preceded by a pre-emphasis network which increases the magnitude of the higher frequency components by an amount proportional to their frequency.

This is then counteracted by a de-emphasis network at the receiver which reverses the operation. Commercial FM therefore is strictly neither FM nor PM but a combination of both. Clearly, the distinction is irrevelant so far as the system is concerned, the only difference between the two being that of the preconditioning of the modulation signal.

Generally, then, an FM or PM modulation system is as illustrated in figure 3.7. The voltage controlled oscillator (VCO) is a sinusoidal oscillator, the output frequency of which is proportional to the input voltage; if the input voltage is f(t) volts, the output frequency is $f_c + f(t)\ \Delta f$ cps.

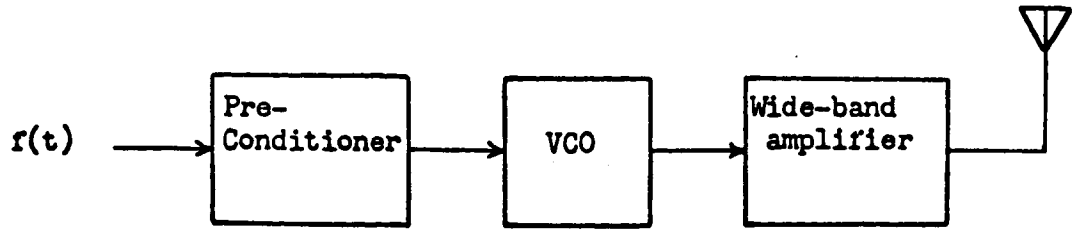


Figure 3.7 An FM Transmitter

There are a number of ways of implementing the block diagram of figure 3.7. However, since we are interested primarily in the system rather than in its particular realization, suffice it to observe that voltage controlled oscillators can be designed to give the desired performance over a wide frequency range.

3.6 Demodulation of Angle Modulated Signals

There are a number of ways in which an FM signal may be demodulated. Any device which is capable of converting a frequency variation into an amplitude variation can serve as an FM demodulator. Such a device is called a frequency discriminator. Suppose, for example, that the FM signal is passed through a filter with the characteristics

$$\left| H(j\,2\pi f) \right| = Kf \qquad\qquad f_c - \Delta f < \left| f \right| < f_c + \Delta f$$

Clearly, the output amplitude is proportional to the input frequency as desired and the FM signal is thereby demodulated. This, in fact, is a somewhat simplified version of a commercial FM discriminator

Another FM demodulator can be designed from the following point of view: Suppose we have, at the receiver, a VCO which is identical to that at the transmitter. If we then make a preliminary estimate of the amplitude of the modulating signal and apply this to the VCO, the similarity between the output of the VCO and the received FM

signal will provide us with a measure of the accuracy of this estimate. We could then use this comparison to improve our original estimate. One way in which this may be accomplished is illustrated in figure 3.8.

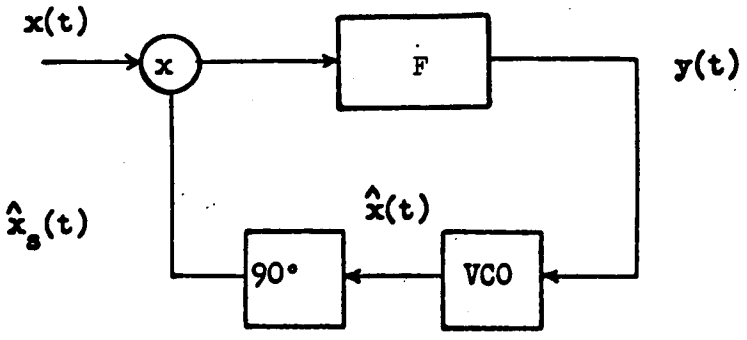


Figure 3.8 A Phase-Locked Loop

This device, called a phase-locked loop, consists of a multiplier, a filter F, a VCO and a device which shifts the phase of the VCO output by 90°. To analyze its behavior suppose that the signal x(t) is

$$x(t) = \sqrt{2}\ \sin(\omega_c t + \theta_1)$$

and suppose that the VCO output is

$$\hat{x}(t) = \sqrt{2}\ \sin(\omega_c t + \theta_2)$$

where $\theta_e = \theta_1 - \theta_2$ represents a small tracking error. Then the product $x(t)\ \hat{x}_s(t)$, where $\hat{x}_s(t)$ represents the shifted version of $\hat{x}(t)$, is formed, yielding

$$x(t)\ \hat{x}_s(t) = 2\ \sin(\omega_c t + \theta_1)\ \sin(\omega_c t + \theta_2 + \frac{\pi}{2})$$

$$= \cos(\theta_e - \frac{\pi}{2}) - \cos(2\omega_c t + \theta_1 + \theta_2 + \frac{\pi}{2})$$

The last term is a high-frequency component and will be eliminated by the combined action of the VCO and the filter F. The low-frequency

component $\cos(\theta_e - \frac{\pi}{2}) = \sin\theta_e \approx \theta_e$ (the last step follows from the assumption that the phase error θ_e is small) is the input to the VCO. Suppose θ_e is positive. Then the VCO frequency is increased to something slightly greater than ω_c thereby decreasing the difference between θ_1 and θ_2 and hence decreasing θ_e. Similarly, if θ_e is negative, the VCO frequency is decreased, again decreasing the absolute value of the difference between θ_1 and θ_2. The loop therefore acts to reduce the phase error to zero.

Now suppose θ_1 varies with time, $\theta_1 = \theta_1(t)$. The loop will again act in such a way as to keep the phase error nearly zero. Then $\theta_2(t) \approx \theta_1(t)$. The difference between the VCO center frequency ω_c and its actual frequency is proportional to its voltage input. Since the instantaneous frequency of the VCO output is $\frac{d}{dt}\left(\omega_c t + \theta_2(t)\right) = \omega_c + \frac{d\theta_{2(t)}}{dt}$ the input to the VCO must have amplitude $k\frac{d\theta_{2(t)}}{dt} \approx k\frac{d\theta_1(t)}{dt}$ where k is a constant of proportionality. Consequently, if the input to the loop is a frequency modulated signal, $\theta_1(t) = \Delta\omega \int f(t)\,dt + \theta_o$, then the input to the VCO is just

$$y(t) \approx k\frac{d\theta_1(t)}{dt} = k\Delta\omega f(t) \qquad (3.23)$$

and the desired signal is recovered.

3.7. Angle Modulation Noise Analysis

To determine the effect of noise at the input let $x(t) = n(t)$ be white noise. Then $n(t)\,\hat{x}_s(t) = n(t)\sqrt{2}\sin\left[\omega_c t + \theta_2(t) + \frac{\pi}{2}\right] = n_1(t)$ is also white noise, as was observed in section 3.4, and with the same power spectral density N_0 as $n(t)$.

Now consider the situation in which $x(t) = \sqrt{2}\, A \sin(\omega_c t + \theta_1(t)) + n(t)$ and $\hat{x}_s(t) = \sqrt{2} \sin(\omega_c t + \theta_2(t) + \frac{\pi}{2})$ where again, it is assumed $\theta_1(t) - \theta_2(t)$ is small. The low frequency term of the product $x(t)\hat{x}_s(t)$ is given by

$$\left[x(t)\, \hat{x}_s(t)\right] = A \sin\left[\theta_1(t) - \theta_2(t)\right] + n_1(t)$$

$$\approx A\left[\theta_1(t) - \theta_2(t)\right] + n_1(t) \,. \tag{3.24}$$

Since $\theta_2(t)$ is adjusted by the action of the loop to keep the error signal and hence the input to F small, it follows that

$$\theta_2(t) \approx \theta_1(t) + \frac{n_1(t)}{A} \tag{3.25}$$

and, consequently,

$$y(t) = k \frac{d\theta_2(t)}{dt} \approx k \frac{d\theta_1(t)}{dt} + \frac{k}{A} \frac{dn_1(t)}{dt}$$

$$= k\, \Delta\omega f(t) + \frac{k}{A} \frac{dn_1(t)}{dt} \tag{3.26}$$

Since the desired output is $k\, \Delta\omega\, f(t)$, the term $\frac{k}{A} \frac{dn_1(t)}{dt}$ represents output noise.

To determine the effect on the noise of taking its derivative, suppose that it consists of a single frequency component, $n(t) = a_n \sin(\omega_n t + \theta_n)$ Then the derivative of the noise consists of the component

$$a_n \omega_n \cos(\omega_n t + \theta_n) \tag{3.27}$$

and has a magnitude equal to ω_n times the magnitude of the original noise. In general, if the power spectral density of the noise is N_0, the power spectral density of its derivative is just $(2\pi f)^2 N_0$.

The signal-to-noise ratio at the output of the FM demodulator is determined as follows: The power in the received signal $k\Delta\omega f(t)$ is, of course,

$$S = k^2(\Delta\omega)^2 P_f \tag{3.28}$$

where P_f is the power in the modulating signal. Since the noise spectral density is $(2\pi f)^2 N_0$, the total noise power is

$$N = \frac{k^2 N_0}{A^2} \int_0^W (2\pi f)^2 df = (2\pi)^2 \frac{k^2 N_0 W^3}{3A^2} \tag{3.29}$$

where W is the bandwidth of the output signal. Evidently, W must be on the order of f_M, the maximum frequency component of the modulating signal, since no higher frequencies are on interest. (If the loop itself did not eliminate all frequencies greater than f_M cps, it could be followed by a low-pass filter which did.) Therefore the output signal-to-noise ratio is, approximately,

$$\left(\frac{S}{N}\right)_{FM} = \frac{3A^2(\Delta\omega)^2 P_f}{(2\pi)^2 N_0 f_M^3} = 3\left(\frac{\Delta\omega}{\omega_M}\right)^2 \frac{A^2 P_f}{N_0 f_M}$$

$$= 3\beta_M^2 \left(\frac{S}{N}\right)_{SSB} \tag{3.30}$$

Recalling that, for large modulation indices, β_M may be interpreted as the ratio of the bandwidth needed with FM to that necessary for conventional AM or DSB transmission, we see that

the signal-to-noise ratio improvement in FM is proportional to the square of this bandwidth multiplication factor β_{FM}. Consequently, FM provides a means for increasing the bandwidth to obtain improved performance. Since increasing the FM bandwidth by β achieves the same results as increasing the signal power by β^2, FM may also be regarded as a method of exchanging power for bandwidth to keep the same performance.

The analysis of the signal-to-noise performances of PM follows along the same lines as that for FM, except that rather than the signal y(t) of figure 3.10, we are now interested in its integral. That is, since $\theta_2(t) \approx \theta_1(t) + \frac{n_1(t)}{A}$ where, in this case, $\theta_1(t) = \Delta\Omega f(t)$, it follows that $\theta_2(t)$ is the quantity of interest, not its derivative y(t). Thus the input to the VCO must also be passed through an integrator in order to yield the desired output. The output signal power is $(\Delta\theta)^2 P_f$ while the noise power is $\frac{1}{A^2} N_0 f_M$ resulting in a signal-to-noise ratio

$$\left(\frac{S}{N}\right)_{PM} = (\Delta\Omega)^2 \frac{A^2 P_f}{N_0 f_M} = (\Delta\Omega)^2 \left(\frac{S}{N}\right)_{SSB} \qquad (3.31)$$

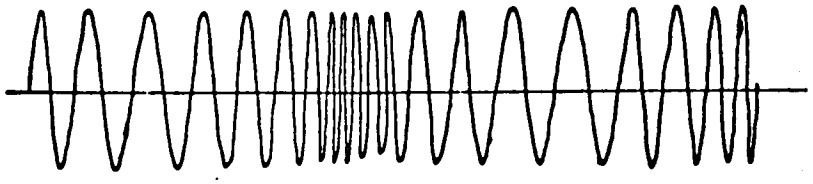
In discussing phase-locked loop demodulation of FM (and PM) we have made some assumptions which should be emphasized. In particular, the conditions were assumed to be such that the VCO phase output was sufficiently close to the input phase to enable us to make the approximation $\sin\ \theta_1(t) - \theta_2(t) \simeq \theta_1(t) - \theta_2(t)$. However, the loop dynamics require that $\theta_2(t)$ $\theta_1(t) + n_1(t)/A$. Clearly, if the term $n_1(t)/A$ represents a phase angle of more than, say, 10°, then this approximation becomes unacceptable. But since the power of the normalized noise $n_1(t)/A$ within the bandwidth f_M of the signal

is $\frac{N_0 f_M}{A^2}$, it follows that $\frac{n_1(t)}{A}$ will be small compared to 10° only so long as the term $\frac{N_0 f_M}{A^2}$ is sufficiently small. As $\frac{N_0 f_M}{A^2}$ becomes large, it will not infrequently happen that this noise term causes $\theta_2(t)$ to be enough different from $\theta_1(t)$ so that the loop can no longer track the input. Generally, then, it is necessary to require that

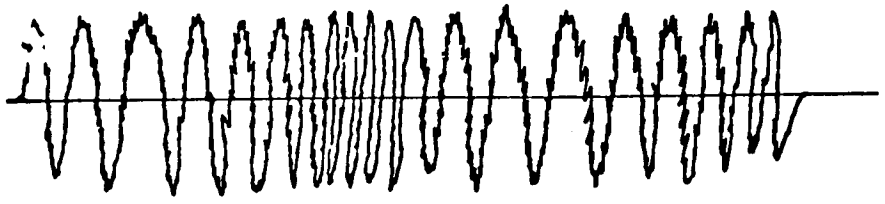
$$\frac{A^2}{N_0 f_M} \geq 36 \tag{3.32}$$

in order that all the assumptions made are reasonable. While the demodulated signal may be meaningful for smaller values of $A^2/N_0 f_M$ than this, the performance rapidly deteriorates as this ratio is further decreased.

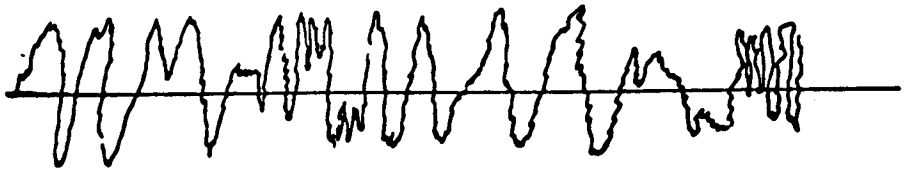
This threshold effect when the noise becomes sufficiently large, is characteristic of any FM demodulating scheme. This can be seen intuitively by referring to figure 3.9 . A frequency modulated signal is shown in figure 3.9 (a). The same signal is shown in figure 3.9 (b) as perturbed by a small amount of additive noise, and in figure 3.9 (c) as altered by noise with a considerably greater power. Since the information is conveyed in an FM signal by the instantaneous frequency, a measure of the effect of the noise exists in the comparison of the position of the zero-crossings before and after the addition of the noise. It is seen that, as the noise increases, some zero crossings will be added by the noise while others will be eliminated entirely. When the noise reaches a level at which this phenomena becomes relatively common, the demodulated signal rapidly deteriorates.



(a) A frequency modulated wave ....



(b)... with small additive noise



(c)... with large additive noise

Figure 3.11

We have not commented on the characteristics of the filter F. Clearly, it is to be chosen if possible so that $\theta_2(t)$ $\theta_1(t)$ regardless of the variation of $\theta_1(t)$, even in the presence of the noise $n_1(t)$. Techniques are available for mathematically specifying the optimum filter when the signal and noise spectra are known. Nevertheless, if the normalized noise power $N_0 f_M/A^2$ is large, the difficulties mentioned above remain, regardless of the filter used.

3.8 Some Applications of Phase-Locked Loops

Before concluding this Section it is well to remark that phase-locked loops have many applications other than FM or PM demodulation. In particular, in the case of DSB and SSB modulation the suppressed carrier could be tracked by such a device, and thereby reproduced with the accuracy needed for product

demodulation. The analysis of the phase-locked loop in this situation is identical to that presented in the previous section with two exceptions: 1) The phase of the received signal $\theta_1(t)$ is constant except for a small variation caused by instabilities in the transmitter oscillator, movement of the transmitter relative to the receiver, and perhaps by random fluctuations caused by the transmission medium. It is not caused to vary deliberately, and hence the bandwidth of $\theta_1(t)$ is very much less here than in the case of FM demodulation; 2) The desired signal output is not $y(t)$ but rather $\hat{x}(t)$ since it is the carrier itself, not just its phase which is to be estimated. The phase error of the estimate $\hat{x}(t)$, it was seen, is just $n(t)/a$ (where $\sqrt{2}\,a$ is the carrier amplitude) and represents an effective phase error power N_0B_L/a^2 (where B_L is the loop bandwidth.) Thus, since the loop bandwidth B_L can be very narrow the phase error can be reasonably small, even for quite small values of a. Since the phase error power N_0B_L/a^2 is the expected value of the square of the phase error, the square root of this quantity gives an estimate of the magnitude of the phase error which is encountered. By requiring

$$\sqrt{\frac{N_0B_L}{a}}$$

to be less than 1/6 radian, for example, one can be reasonably sure that the phase error remains within tolerable limits. It will be recalled from section 3.4 that

$$\left(\frac{S}{N}\right)_{SSB} = \frac{P_T}{N_0f_M}$$

where P_T is the total power in the received signal, N_0, the noise spectral density, and f_M the signal bandwidth. This was true under the assumption

that the ratio of the power in the carrier to that in the modulation was negligibly small. Suppose the output signal-to-noise ratio $(S/N)_{SSB}$ must be at least unity; i.e., the signal power must be at least as great as the noise power, a generally quite marginal condition. Then

$$\frac{N_0 B_L}{a^2}\left(\frac{S}{N}\right)_{SSB}$$

$$= \frac{N_0 B_L}{a^2}\,\frac{P_T}{N_0 f_M} = \frac{1}{36}$$

and

$$\frac{P_T}{a^2} = \frac{1}{36} \cdot \frac{f_M}{B_L}$$

As an axample, let f_M = 6000 cps. The effective loop bandwidth of the carrier tracking loop can often be made as small as 1.0 cps or less. Thus

$$\frac{P_T}{a^2} = \frac{1}{6} \times 10^3$$

and, indeed, the required carrier power is negligible.

These same comments are equally applicable to any filter of bandwidth B_L. The advantage of the phase-locked loop over ordinary filters is that it is able to track the signal, thereby requiring a much smaller bandwidth. Suppose, for example, that in the course of a transmission the carrier frequency could vary by as much as 20 cps, actually a reasonably small variation when all possible causes, such as doppler shifts and transmitter instabilities are taken into consideration. A conventional filter would need a bandwidth of at least 20 cps in order not to lose the signal entirely. If the rate at which the frequency varied were small, however, the phase-locked loop could very likely track it with a bandwidth of 1 cps or less.

4 PULSE MODULATION

In this chapter we consider modulation techniques which differ fundamentally from those of the previous chapter. As a point of departure we begin with a discussion of the sampling theorem.

4.1 The Sampling Theorem

It is rather obvious that many of the measurements which are monitored on a space craft do not need to be observed continually. Temperatures for example generally vary quite slowly and readings need be taken only every minute or even every hour, certainly not continuously. It is perhaps surprising that not only temperature but any continuous time function can be observed only periodically without any loss of information concerning the original signal provided that the signal power spectrum vanishes for frequencies greater than some finite frequency W. The complete time function can be reconstructed from the periodic samples alone.

To verify this consider the periodic function S(t) illustrated in figure 4.1

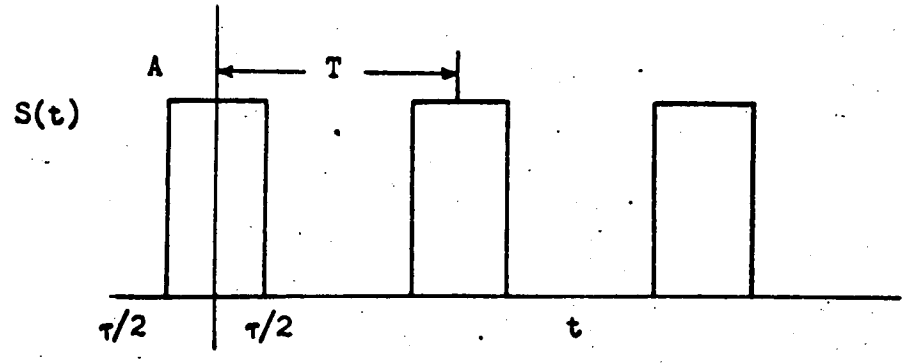


Figure 4.1 The Periodic Function S(t)

Since S(t) is periodic it can be expanded in a Fourier series (cf. section 1.1)

$$S(t) = \frac{2}{T} \sum_{n=1}^{\infty} c_n \cos \omega_n t + \frac{A\tau}{T} \qquad \omega_n = \frac{2\pi n}{T} \tag{4.1}$$

where

$$c_n = A\tau \frac{\sin \omega_n \tau/2}{\omega_n \tau/2} \quad .$$

Now suppose we have some time function x(t) which we wish to transmit. Consider the product x(t) S(t). The situation is identical to that encountered in Section 3.2. If x(t) has the spectrum shown in Figure 4.2(a), the spectrum of the product is as depicted in Figure 4.2(b) provided that W < 1/2T

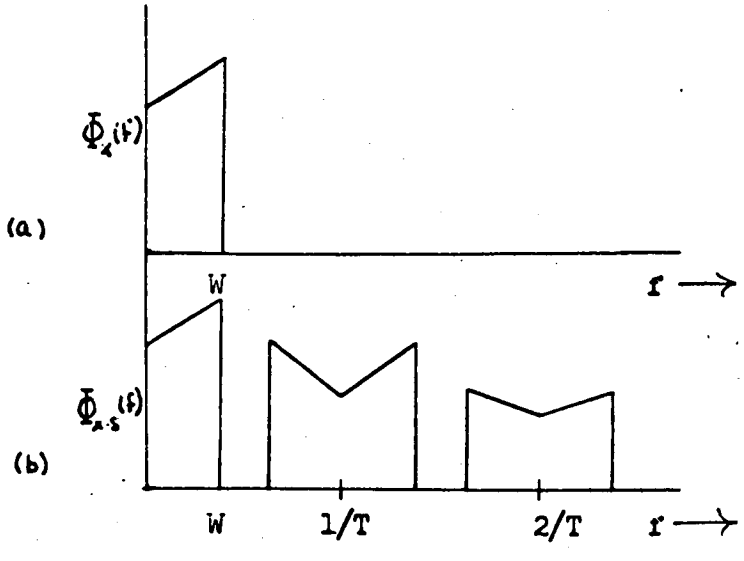


Figure 4.2 Power Spectra of x(t) and x(t)S(t).

Thus, the higher components can be filtered off leaving just the original process $x(t)$ multiplied by the constant $\frac{A\tau}{T}$. Consequently, the time function $x(t)S(t)$ contains all of the information of the signal $x(t)$. Note that this is true regardless of the value of τ so long as $\frac{A\tau}{T} > 0$. In particular if we let $\tau \to 0$ and $A \to \infty$ such that $\frac{A\tau}{T} = 1$, $S(t)$ becomes a periodic sequence of infinitely narrow pulses and $x(t)$ is observed only at the instants of time $t = iT$, for all integers i. All of the information needed to reconstruct the signal $x(t)$, when the spectrum of $x(t)$ is limited to W cps, is contained in the values of the amplitude of $x(t)$ at the instants of time $x(nT)$ subject only to the constraint that $T < \frac{1}{2W}$.

This result is known as the sampling theorem. It states that we need only concern ourselves with periodic samples of a bandwidth limited time function. Only the sequence of numerical values $x(nT)$ need be transmitted. The complete function $x(t)$ can be reconstructed at the receiver by generating a series of very narrow pulses of area $x(nT)$ and passing them through a lowpass filter.

4.2 Time and Frequency Multiplexing

There are a number of advantages associated with sampled data telemetry systems. First of all, as we shall see, some rather elegant techniques have been devised for transmitting sampled data, methods at the same time relatively easily implemented and achieving large signal-to-noise ratio increases at the expense of bandwidth.

In addition, it is often considerably easier and more efficient to handle sampled data than continuous data. Typically, a spacecraft may contain 100 to 1000 data sources. Some method must

be used to keep the information from each source separate. One method for doing this, called frequency multiplexing, consists of forming the products

$$x_i(t) \sin \omega_i(t)$$

for each data signal $x_i(t)$, $i = 1, 2, \ldots,$. The frequencies $\frac{\omega_i}{2\pi}$ must be such that the spectra of each of the signals do not overlap as shown in figure 4.3 . The signal

$$x(t) = \sum_{i=1}^{N} x_i(t) \sin \omega_i(t) \qquad (4.2)$$

then has the composite spectrum of figure 4.3, and $x(t)$ may be treated as a single source with a bandwidth

$$W = W_o + 2\sum_{i=1}^{N} W_i \ , \qquad (4.3)$$

where W_i is the bandwidth of the process $x_i(t)$.

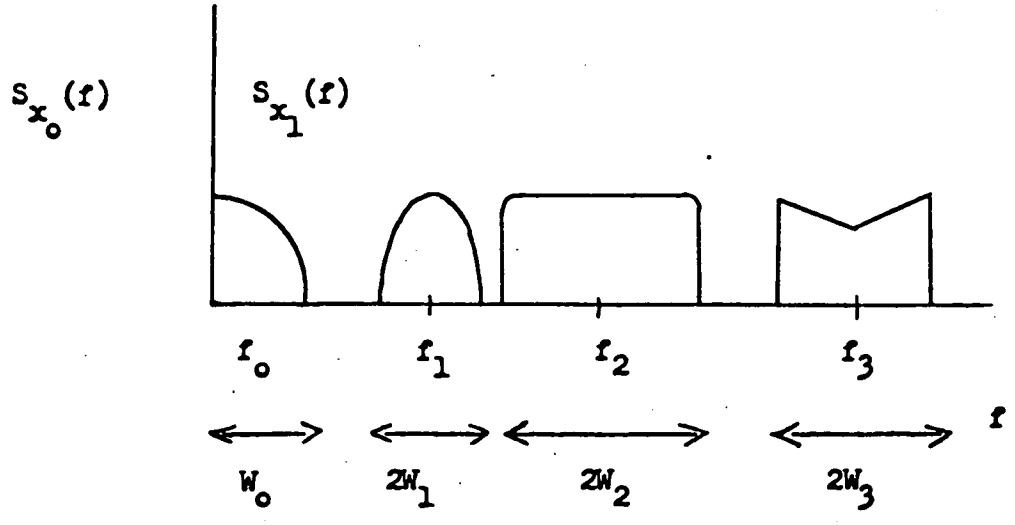


Figure 4.3 Frequency Multiplexing

Since the individual spectra do not overlap, the different signals $x_i(t)$ can be reconstructed at the receiver by proper filtering. Unfortunately, this method

demands that each source be followed by a device for forming the product $x_i(t) \sin \omega_i t$, a procedure which is quite inefficient, particularly on board a spacecraft.

An alternative is to sample each of the signals $x_i(t)$, represent the samples as pulses of duration T/N where T is the sampling rate (here it is assumed that all signals have the same bandwidth so that $T = 1/2W$ is the same for all $x_i(t)$) and <u>time multiplex</u> these samples as shown in Figure 4.4.

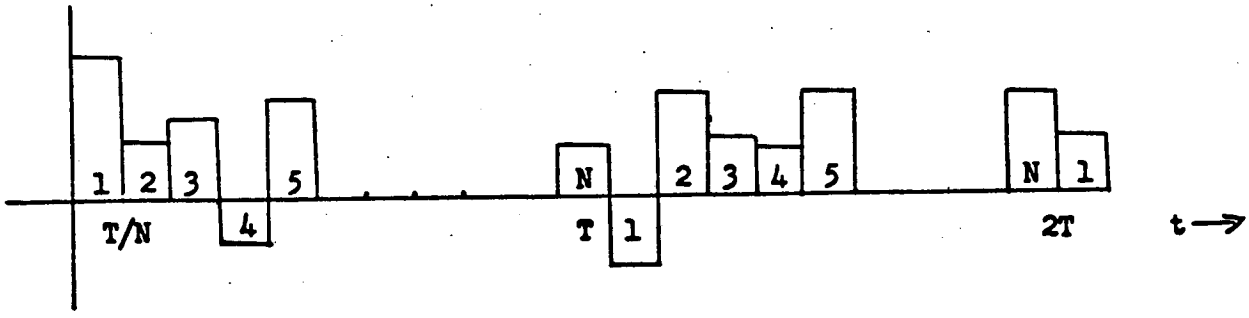


Figure 4.4. Time Multiplexing

The pulse labeled i corresponds to a sample of the process $x_1(t)$. If the bandwidths of the pulses are not the same, the sampling rates must be different, or all rates must be equal to that demanded by the signal with the largest bandwidth. Different sampling rates can readily be accomodated so long as they are integrally related. That is, suppose $x_1(t)$ has a bandwidth which is twice as great as $x_2(t)$ and the two are to be time multiplexed. Since $x_1(t)$ must be sampled twice as often as $x_2(t)$, they can be multiplexed as shown in figure 4.5. Thus, in time T two samples of $x_1(t)$ are transmitted while one sample of $x_2(t)$ is transmitted; both are sampled at the correct rate.

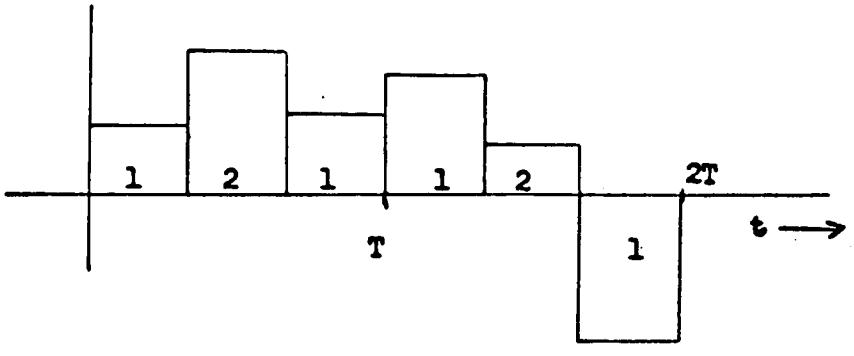


Figure 4.5 Time Multiplexing of Signals with Unequal Bandwidths

A time multiplexing system involves only the problem of commutation or the interspacing of samples from the various sources at the proper rate. No power consuming auxillary equipment is required other than a moderate number of switching devices.

It might be supposed that, since each data signal is only being observed for an infinitesimal fraction of the time, the bandwidth requirements could be considerably reduced. Actually this is not the case, for recall from Section 1.5 that the bandwidth B necessary to transmit a signal which exhibits variations of interest on the order of every τ seconds must satisfy the inequality $B > 1/2\tau$. The sampling theorem states that a signal of bandwidth W must be sampled at least every $T = 1/2W$ seconds. If this sample were transmitted as a pulse, the pulse could last only T seconds before the next sample must be sent. Thus the pulse width cannot be greater than $\tau = T = 1/2W$ and, hence, the bandwidth occupancy must be, at least, $B = 1/2\tau = W$, the bandwidth of the signal.

The same relationship applies to multiplexed signals; frequency multiplexed signals require bandwidths at least as great as the sum of the bandwidths of the separate signals.

$$W = \sum_i W_i \tag{4.4}$$

(Actually, the frequency multiplexed bandwidth defined in Equation (4.3) is about twice this value. However, it will be recalled from Section 3.3 that the bandwidth can be halved without any loss of information. Thus, in the case of single-sideband frequency multiplexing the two multiplexing schemes yield the same bandwidth. Single-sideband multiplexing could be accomplished, of course, only at the expense of additional equipment.) A time multiplexed system involving N data sources has only T/N seconds per pulse as was seen in Figure 4.4. Thus, the bandwidth is increased by a factor of N in the case of equal signal bandwidths $W_i = W$; ($\tau = 1/2NW$, $B = NW$). Similarly, the frequency

multiplexed signal bandwidth is increased by the same factor under the same condition as seen from Equation (4.4). If the bandwidths are not equal, the frequency multiplexed signals can be spaced more efficiently, in general, since there is no necessity that the sampling rates be integrally related. However this advantage is offset by the necessity of single-sideband multiplexing to avoid increasing the bandwidth by a factor of two.

4.3 Pulse Modulation Systems and Matched Filtering

In the sections that follow a number of pulse modulation methods will be discussed. In order to simplify the discussion it will be assumed that the data input to the transmitter consists of a sequence of samples at some average rate, say R samples per second. Thus each sample has $T = 1/R$ seconds in which to be transmitted. It is unimportant whether this sequence comes from one source or is the time multiplexed output from a number of sources. A pulse modulation system involves the transmission of a particular waveform $f(t)$ representing the sample in question for a period of time T seconds. The transmitted signal, therefore, is allowed to change form only every T seconds.

Before proceeding to discuss various pulse modulating schemes in more detail, it is of interest to consider the generic form of the demodulators for pulse modulation. The pulse modulated signal, as observed, is characterized by a waveform $f_i(t)$ which is transmitted without interruption for the time interval $\nu T < t < (\nu+1)T$.

There may be a finite or an infinite number of waveforms $f_i(t)$ which can be transmitted during each interval. (If the signal can assume a continuum of amplitudes, for example, there must be infinite number of waveforms to represent every possible signal. The set of waveforms $f_i(t)$ which can be transmitted is known at the receiver. The received signal will be perturbed by noise so that it will generally not match exactly any of the signals in this set. Suppose the signal $f_j(t)$ is transmitted and the signal $y(t) = Af_j(t) + n(t)$ is received where A is assumed to be a known constant. Hopefully $y(t)$ will resemble $Af_j(t)$ more closely than it resembles any of the other signals in the set $Af_i(t)$. If not, there is little chance of making a correct decision as to the transmitted signal. An ideal receiver would then consider the entire set $Af_i(t)$ and determine that waveform which most nearly resembles the received signal. But how do we measure how closely one waveform resembles another? One method might be to determine the average value of the difference between the received waveform and each of the possible waveforms

$$\epsilon_j = \frac{1}{T} \int_{vT}^{(v+1)T} (y(t) - Af_j(t)) \, dt$$

and estimate the received signal as that corresponding to the ϵ_j with minimum absolute value. This would not be particularly effective, of course, since $y(t)$ could be much greater than $Af_j(t)$ over part of the interval, and much less over the remainder of the interval, and still yield a value of $\epsilon_j \approx 0$. An approach which avoids this difficult (and, in fact, one which can be shown to be optimum

for many situations, including that normally encountered in space communications) is to determine the mean squared error:

$$\epsilon_j^2 = \frac{1}{T} \int_{\nu T}^{(\nu+1)T} (y(t) - Af_j(t))^2 \, dt. \tag{4.5}$$

If $\min_j \epsilon_j^2 = \epsilon_k^2$, the receiver concludes that the received signal was $f_k(t)$.

Note that

$$\epsilon_j^2 = \frac{1}{T} \int_{\nu T}^{(\nu+1)T} y^2(t) \, dt - \frac{2A}{T} \int_{\nu T}^{(\nu+1)T} y(t) f_j(t) \, dt + \frac{A^2}{T} \int_{\nu T}^{(\nu+1)T} f_j^2(t) \, dt$$

and, since the first term on the right is independent of j, the optimum signal is that for which the quantity

$$\frac{1}{T} \int_{\nu T}^{(\nu+1)T} y(t) f_j(t) \, dt - \frac{A E_j}{2T} \tag{4.6}$$

is a maximum. Here E_j represents the energy in the jth signal $f_j(t)$. The term

$$g_j\left[(\nu+1)T\right] = \frac{1}{T} \int_{\nu T}^{(\nu+1)T} y(t) f_j(t) \, dt \tag{4.7}$$

is called the cross-correlation between the signals $y(t)$ and $f_j(t)$. Consequently,

such a receiver is frequently called a correlation receiver. The process is, of course, repeated for all integer values of ν, so long as signals are being received.

4.4 Pulse Amplitude Modulation (PAM)

Perhaps the most obvious method for transmitting sampled data is pulse amplitude modulation. If the data sample is x_ν, the signal $\sqrt{2x_\nu}\, B \sin \omega_c t$ is transmitted over the internal $\nu T < t < (\nu + 1)T$, the received signal then becoming $y(t) = \sqrt{2A}\, x_\nu \sin \omega_c t + n(t)$.* As discussed in the previous section, the optimum detector forms the quantity

$$g_i\left[(\nu+1)T\right] - \frac{A}{2T} E_i$$

$$= \frac{1}{T} \int_{\nu T}^{(\nu+1)T} y(t) \sqrt{2}\, x_i \sin \omega_c t \, dt - \frac{A}{2T} \int_{\nu T}^{(\nu+1)T} 2x_i^2 \sin^2 \omega_c t \, dt \tag{4.8}$$

for every possible amplitude x_i of the received signal, and selects the largest of those as the best estimate of the received signal. But a condition that the quantity $g_i\left[(\nu+1)T\right] - \frac{A}{2T} E_i$ be a maximum is that

*There will be, in general, a phase and probably even a frequency shift between the transmitter and the receiver. However, this will cause no difficulty so long as the received phase and frequency are determined at the receiver and used to generate the local signals $f_i(t)$. This knowledge will be assumed here so that the phase and frequency shift can safely be ignored.

$$\frac{d}{dx_i}\left\{ g_i\left[(\nu+1)T\right] - \frac{A}{2T}\ E_i \right\} = 0$$

or that $$\frac{1}{T}\int_{\nu T}^{(\nu+1)T} y(t)\sqrt{2}\sin\omega_c t\ dt = \frac{x_i A}{T}\int_{\nu T}^{(\nu+1)T} 2\sin^2\omega_c t\ dt \qquad (4.9)$$

$$= x_i A$$

where the carrier frequency f_c has been chosen to be some multiple of half the reciprocal of the pulse period T, $f_c = \frac{\omega_c}{2\pi} = \frac{k}{2T}$ for some integer k. The optimum estimate $\hat{x}_\nu$ of the amplitude of the received signal, then is

$$\hat{x}_\nu = \frac{1}{A\,T}\int_{\nu T}^{(\nu+1)T} y(t)\sqrt{2}\sin\omega_c t\ dt \qquad (4.10)$$

and the receiver is simply that illustrated schematically in figure 4.6.

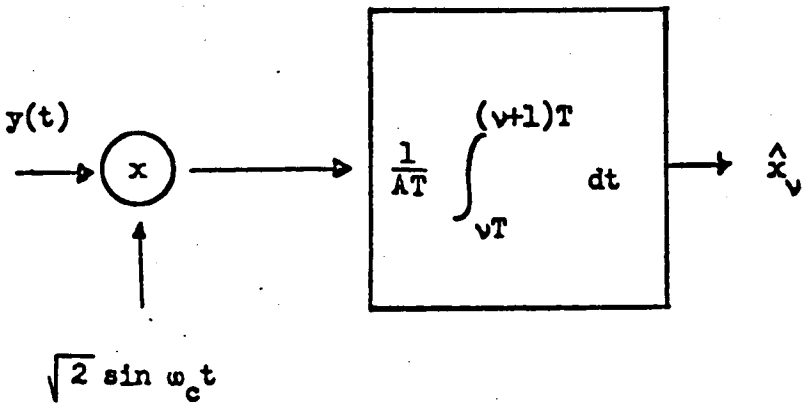


Figure 4.6. A PAM Demodulator

To determine the signal-to-noise ratio at the output of a PAM demodulator we note that the output signal can be written

$$\hat{x}_\nu = \frac{2}{T}\int_{\nu T}^{(\nu+1)T} x_\nu \sin^2 \omega_c t \; dt + \frac{1}{AT}\int_{\nu T}^{(\nu+1)T} n(t)\sqrt{2}\sin \omega_c t \; dt$$

$$= x_\nu + n_\nu \qquad (4.11)$$

where x_ν is the output due to the received signal, and n_ν that due to the noise. The output signal power therefore is just the power P_x in the modulating signal x_ν (which is presumably equal to the power in the original signal x(t)). As we have noted several times, multiplying white noise by a sinusoid of amplitude $\sqrt{2}$ does not alter its whiteness nor its spectral density. It can be shown that the integrator acts as a filter with a noise bandwidth

$B_n = \frac{1}{2T}$ so that the output noise power is $\frac{N_0}{2T\,A^2}$ (the A^2 term due to the fact that the noise has been divided by A, the noise power therefore is reduced by the factor A^2). The output signal-to-noise ratio is consequently,

$$\left(\frac{S}{N}\right)_{PAM} = \frac{A^2 P_x}{N_0 \left(\frac{1}{2T}\right)} = \frac{A^2 P_x}{N_0 B_n} \tag{4.12}$$

An interesting measure of the effective bandwidth of a signal is afforded by asking the question: How far in frequency must different communication channels be separated if the cross-modulation between any two of them is to be kept to an insignificant level? Suppose, in fact, that a number of PAM channels were to be operated simultaneously at the carrier frequencies ω_i, $i = 1, 2, \ldots$. Then the effect of the signal in the i^{th} channel on the output of the j^{th} channel demodulator is simply

$$\frac{1}{AT} \int_{\nu T}^{(\nu+1)T} y_i(t) \sqrt{2} \sin \omega_j t \; dt$$

$$= \frac{1}{AT} \int_{\nu T}^{(\nu+1)T} x_\nu(i)\, 2 \sin \omega_i t \sin \omega_j t \; dt \tag{4.13}$$

$$= \frac{x_\nu(i)}{AT} \left[\int_{\nu T}^{(\nu+1)T} \cos (\omega_i - \omega_j)t \; dt + \int_{\nu T}^{(\nu+1)T} \cos (\omega_i + \omega_j)t \; dt \right]$$

which is identically zero if ω_i and ω_j are different non-zero integer multiples of the term $\frac{\pi}{T}$. Thus, in order to keep the cross-modulation zero it is necessary to separate the channels in frequency by an amount $f_i - f_j = \frac{k}{2T}$, for any value of $k = 1, 2, \ldots$. The effective bandwidth of each channel is therefore $B_{eff} = \frac{1}{2T} = W$ cps where W is the bandwidth of the sampled signal (see section 4.2). Consequently, Equation (4.12) can be rewritten in terms of the effective bandwidth to yield

$$\left(\frac{S}{N}\right)_{PAM} = \frac{A^2 P_x}{N_0 B_{eff}} \qquad (4.14)$$

Note that this is exactly the same relationship that was obtained for Suppressed Carrier amplitude modulation.

4.5 Phase-Shift Keyed Modulation (PSK)

Another rather common pulse modulation technique, called phase-shift keying, is to transmit the signal

$$\sqrt{2} \quad \sin(\omega_c t + x_\nu) \qquad \nu T < t < (\nu+1)T \qquad (4.15)$$

to convey the data x_ν. Here, of course, $0 < x_\nu < 2\pi$ in order that there be no ambiguity at the receiver. Thus the phase, rather than the amplitude, conveys the information in a PSK system. The advantage of this method over PAM modulation rests in the fact that the amplitude of the signal remains constant. This is not an insignificant advantage in space telemetry since transmitters which work at a constant amplitude are considerably more efficient than those which must produce variable amplitudes.

From section 4.3, the optimum PSK receiver, since the received signal energy is now independent of x_ν, must form the integrals

$$\int_{\nu T}^{(\nu+1)T} y(t)\sqrt{2}\ \sin(\omega_c t + x_i)dt \tag{4.16}$$

for all $0 < x_i < 2\pi$ and select the largest. But this expression may be written

$$\cos x_i \int_{\nu T}^{(\nu+1)T} y(t)\sqrt{2}\sin\omega_c t\,dt + \sin x_i \int_{\nu T}^{(\nu+1)T} y(t)\sqrt{2}\cos\omega_c t\,dt \tag{4.17}$$

$$= X\cos x_i + Y\sin x_i\ .$$

where $X = \int_{\nu T}^{(\nu+1)T} \sqrt{2}\, y(t)\sin\omega_c t\,dt$ and $Y = \int_{\nu T}^{(\nu+1)T} \sqrt{2}\, y(t)\cos\omega_c t\,dt$

The maximum of these with respect to x_i must satisfy the condition that

$$\frac{d}{dx_i}\left\{X\cos x_i + Y\sin x_i\right\} = 0$$

or that

$$x_i = \tan^{-1}\left(\frac{Y}{X}\right) \tag{4.18}$$

Thus the optimum estimate of x_ν is $\hat{x}_\nu = \tan^{-1}\ \frac{Y}{X}$ and the optimum receiver is that depicted in figure 4.7.

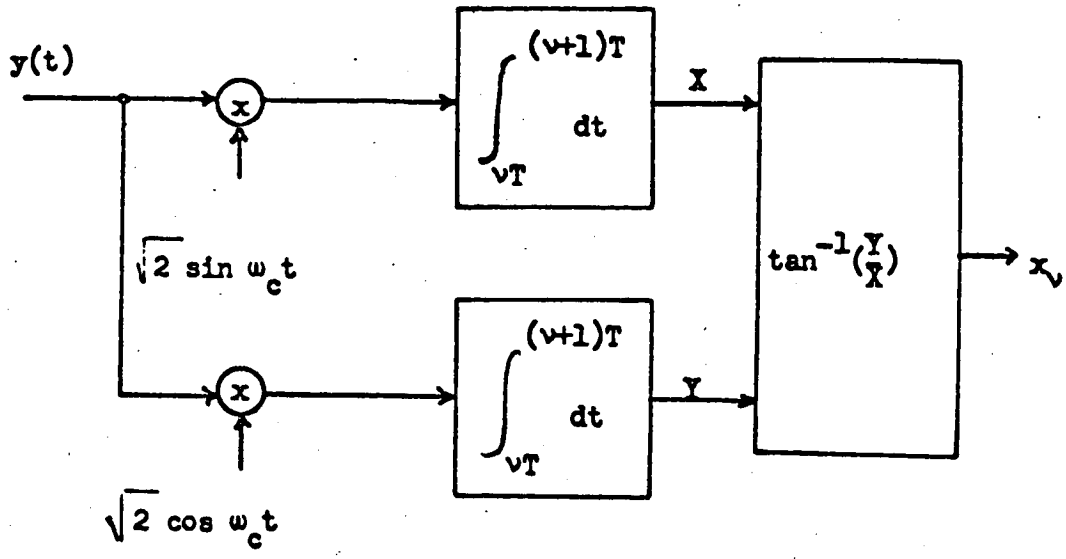


Figure 4.7. A PSK Demodulator

It is observed that in the absence of noise,

$$y(t) = \sqrt{2}\, A \sin(\omega_c t + x_\nu),$$

and

$$X = \int_{\nu T}^{(\nu+1)T} 2A \sin \omega_c t \sin(\omega_c t + x_\nu)\, dt$$

$$= A \int_{\nu T}^{(\nu+1)T} \left[\cos x_\nu - \cos(2\omega_c t + x_\nu)\right] dt$$

$$= AT \cos x_\nu$$

while

$$Y = \int_{\nu T}^{(\nu+1)T} 2A \cos \omega_c t \sin (\omega_c t + x_\nu)\, dt$$

$$= A \int_{\nu T}^{(\nu+1)T} \left[\sin x_\nu + \sin(2\omega_c t + x_\nu)\right] dt$$

$$= A\,T \sin x_\nu$$

where it is again assumed that $\omega_c = \frac{\pi k}{T}$ for some integer k.

Hence

$$\hat{x}_\nu = \tan^{-1} \frac{AT \sin x_\nu}{AT \cos x_\nu} = x_\nu$$

and the estimate of the signal is exact in the absence of noise. The analysis of the output signal-to-noise ratio for PSK is somewhat more involved than that for PAM and will not be carried out here. The results of such an analysis, however, would indicate performance approximately equal to that of PAM. In addition, it is easily verified that essentially the same comments concerning the spectrum of a PAM signal as well as its effective bandwidth occupancy apply equally to a PSK modulated signal. (The effective bandwidth occupancy of a PSK signal is actually twice that of a PAM signal).

4.6 Pulse Code Modulation (PCM)

In all of the modulation schemes discussed up to now, it has been possible to transmit any of a continuum of amplitudes (generally assumed to be bounded by finite values). But several observations strongly suggest that this is not

necessary. In the first place, regardless of the use to which the receiver of the information intends to apply it, he can never use more than a finite number of significant digits for each sample. If for no other reason, the accuracy of the measuring device is always limited to some degree. Moreover, the noise encountered at the receiver insures that the data will contain some inaccuracies nullifying the meaningfulness of all but the most significant digits in the received samples. Thus, so far as the receiver is concerned, it is of little consequence whether or not the data is quantized in amplitude as well as in time. Because many of the measurements are inherently digital in nature anyway (the outputs of counters, for example) it may be of some advantage to quantize all the information so that it can all be processed uniformly on the spacecraft. But is there any advantage so far as the communication system itself is concerned, to quantizing all analogue signals before transmission? The answer is emphatically yes, as the following analysis of a pulse-code modulation system will illustrate.

Since all the data is to be quantized, the communication system need be concerned with only a finite number of signals $f_i(t)$. In particular, suppose the quantized information is represented in binary form. If the output of a device were quantized to four levels, for example, the possible outputs could be represented by the four two-digit numbers

$$\begin{array}{c} 0\ 0 \\ 0\ 1 \\ 1\ 0 \\ 1\ 1 \end{array} \qquad (4.19)$$

In general, if the output is quantized to M levels, each possible output can be represented by $\log_2 M$ (or the smallest integer greater than $\log_2 M$ if it is not an integer) binary symbols or bits. (Of course, bases other than the binary base

could be used, but binary systems possess certain distinct advantages and are by far the most commonly encountered.)

A pulse-code modulation (PCM) system is one in which each of the binary digits used to represent a data sample is individually transmitted as one of two possible waveforms. That is, if the bit is a one, some signal f(t) is transmitted for a duration of T_B seconds; if it is a zero some other signal g(t) is transmitted for the same duration. Then the next bit is similarly transmitted, etc. Typically, g(t) = - f(t). The PCM correlation receiver is shown in figure 4.8. Note that if g(t) is equal

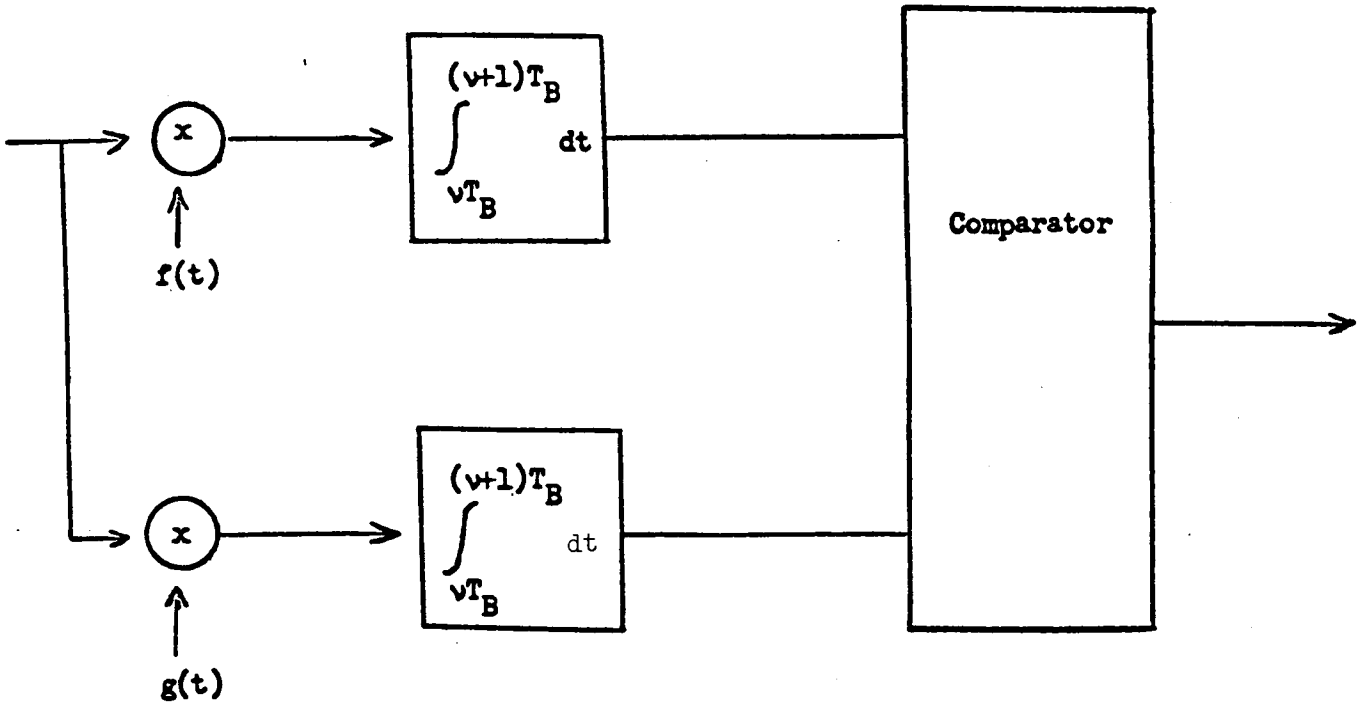


Figure 4.8 An Optimum PCM Detector

to -f(t) only one integrator is needed. The comparator then decides that a <u>one</u> was received if the integrator output is positive, and that a <u>zero</u> was received

if it is negative.

To analyze the output signal-to-noise ratio using this modulation method, let us assume the ratio of the received signal power to the noise power to be sufficiently large to render it is most unlikely that f(t) can be mistaken for g(t) at the receiver, or conversely. The implications of this assumption will be investigated shortly. In this case the noise is due exclusively to the quantization of the signal. That is, if the sampled signal is $s(t) = s(\nu T)$, $\nu T < t < (\nu + 1)T$ the quantized signal $s_q(\nu T)$ is transmitted. Since no mistake is made at the receiver in identifying each of the digits representing $s_q(t)$, this signal is recovered exactly. Thus, the mean-squared noise at the receiver is just

$$\frac{1}{T}\int_{\nu T}^{(\nu+1)T} (s(t) - s_q(t))^2 \, dt = (s(\nu T) - s_q(\nu T))^2 \tag{4.20}$$

Suppose the signal is equally likely to assume any amplitude from -a to a, and that the quantization is such that

$$s_q(\nu T) = \frac{2a(i+\frac{1}{2})}{M}$$

when

$$\frac{2ai}{M} < s(\nu T) < \frac{2a(i+1)}{M} \qquad i = -\frac{M}{2}, -\frac{M}{2}+1, \ldots, \frac{M}{2}-1$$

(4.21)

Then, while the signal varies between the extremes of -a and a and is equally likely to have any value in this range at any given time, the quantization error varies between -a/M and a/M and is equally likely to assume any value in that range. Accordingly, the output signal-to-noise ratio is, from Section 1.6,

$$\left(\frac{S}{N}\right)_{PCM} = \frac{\lim_{T\to\infty} \int_{-T/2}^{T/2} S^2(t)\, dt}{\lim_{T\to\infty} \int_{-T/2}^{T/2} \left(S(t) - S_q(t)\right)^2 dt} = M^2 \qquad (4.22)$$

This suggests that by quantizing to an infinite number of levels, we could obtain an infinite output signal-to-noise ratio. Remember, however, that we have assumed that f(t) and g(t) could be distinguished correctly with very high reliability at the receiver. If signals are to be transmitted every T_0 seconds, and each one is to be quantized into $M = 2^m$ levels then each bit has only

$$T_B = \frac{T_0}{\log_2 M} = \frac{T_0}{m} \qquad (4.23)$$

seconds in which to be transmitted. The performances of the previous pulse modulation systems have shown a strong dependence on the amount of time spent in transmitting any particular waveform. This dependence is equally true here. If, for example, g(t) = -f(t), and the average power in the signal f(t) at the receiver is A^2 and the noise spectral density is N_0, then it is usually necessary to require that

$$\frac{A^2 T_B}{N_0} \geq 5 \qquad (4.24)$$

if the two signals f(t) and -f(t) are to be distinguished reliably enough at the receiver to justify the assumption made in this derivation.

Again, as in the case of PAM, the effective bandwidth necessary to transmit a pulse of time duration T_B is

$$B_{eff} = \frac{1}{2T_B} \tag{4.25}$$

Thus, condition 24 becomes

$$\frac{A^2}{N_0 B_{eff}} \geq 10 \tag{4.26}$$

and is consequently analogous to the conditon of Equation (3.32) in Section Three for the FM threshold.

While PCM exhibits a threshold effect comparable to that of FM, it also demonstrates the same ability to increase communication reliability at the expense of bandwidth. For, from Equations (4.22) and (4.23) we have

$$\left(\frac{S}{N}\right)_{PCM} = M^2 = 2^{2\left(\frac{T_0}{T_B}\right)} \tag{4.27}$$

But $1/2T_0$ is just the effective bandwidth of the signal (cf. Section 4.2) and, in fact, represents the bandwidth necessary to send the sampled signal as a PAM (or SSB) signal (see Section 4.4). Consequently, the bandwidth expansion factor of PCM is

$$\beta_{PCM} = \frac{2T_0}{2T_B} \tag{4.28}$$

so that

$$\left(\frac{S}{N}\right)_{PCM} = 2^{2\beta_{PCM}} \tag{4.29}$$

The reliability increases exponentially with the bandwidth expansion rather than as the square. The trade-off is therefore much more advantageous in the case of PCM than with any of the other systems discussed up to now. The limitation to this increased fidelity, as expressed in equation 26, however, must be observed.

4.7 Coding and Waveform Selection

The last section demonstrated the superior performance of the quantized PCM system. This performance was predicated upon the small likelihood of the erroneous reception of a bit. Thus, the number of quantization levels, and hence the quantization error, is limited by the fact that T_B cannot be made too small or the bit error probability becomes too great. Several techniques are available for decreasing the significance of symbol errors, and hence increasing the number of possible quantization levels. One of these is the use of error-correcting and error-detecting codes.

The subject of coding has received a large amount of attention in the literature in the past decade and is considerably too complex to be covered in any detail here. The principle of coding, however, is quite simple. Since one symbol error occurs relatively infrequently, two symbol errors will occur even more rarely, three still less often, etc. (Note, a binary digit is generally called a bit if it represents the data non-redundantly; i.e. if for every n bits received it is possible to distinguish 2^n data levels. If this is not true, the binary digits are customarily referred to as symbols.) Thus, if instead of the set of bits (19) the symbols

$$\begin{matrix} 0\,0\,0 \\ 0\,1\,1 \\ 1\,0\,1 \\ 1\,1\,0 \end{matrix} \qquad (4.30)$$

were transmitted, one symbol error can no longer cause an error in reception. If, for example, the sequence 0 0 1 is received an error must have occurred. It is not known which sequence was transmitted, but, at least, the error has been detected. Such a code (i.e., data representation) is called an error-detecting code. If this code is altered further by adding two more symbols to each sequence in the following manner

$$\begin{matrix} 0\,0\,0\,0\,0 \\ 0\,1\,1\,0\,1 \\ 1\,0\,1\,1\,0 \\ 1\,1\,0\,1\,1 \end{matrix} \qquad (4.31)$$

it can be verified that, if only one symbol error is made, the resulting sequence is more like the original sequence than any other. If, for example, 0 0 0 0 0 is transmitted and one error is made so that 0 0 1 0 0 is received, the received sequence differs from the first sequence in only one symbol, from the second in two, from the third in two and from the fourth in five. Since one symbol error is made more likely than two or five, the most logical decision is that the transmitted sequence was 0 0 0 0 0. Then single errors can be corrected and codes which do this are called single-error correcting codes. These codes were achieved by adding redundant symbols to the original information bits. By adding further redundancy higher error-correcting properties can be obtained. The communication system has now been altered to include, between the source and the transmitter, an encoder to add redundancy to the data sequence. Similarly, a decoder must be inserted between the receiver and the user.

This procedure does not necessarily decrease the probability of making an error in identifying the original sample at the receiver, however. This is because, in order to keep the transmission rate the same, three and five symbols must be sent, using the codes (30) and (31) respectively, in the same time that two bits would be sent using the code (19). Thus the time per symbol has dropped by a factor of 2/3 and 2/5 respectively in the two cases. Since, as observed in Section 4.6, the probability of a symbol error is a function of the symbol time, symbol errors are more likely using the codes (30) and (31). Thus while error-detecting and error-correcting codes make symbol errors less costly, they also make them more probable. Whether there is a net gain or loss depends upon the code and upon the system parameters.

A second technique for decreasing the probability of an error at the receiver lies in the judicious selection of waveforms to represent the various signals. Suppose we take a more general view of an encoder as a device for accepting n successive data bits and representing them as one of $N=2^n$ waveforms, each different set of n bits corresponding to a different waveform. The encoder then submits to the transmitter a signal which assumes one of N amplitudes and allows the transmitter T_0 seconds to transmit it. It is desired to send this information as reliably as possible, i.e. to minimize the probability that the receiver interprets its input as something other than that which was transmitted. What waveforms should we use to represent the data? Intuitively, the waveforms should be as dissimilar as possible.

In section 4.3 we introduced the concept of cross-correlation as the measure of similarity between two signals. Consider the <u>normalized cross-correlation coefficient</u> ρ_{12} between two time functions $y_1(t)$ and

$y_2(t)$ where ρ_{12} is defined by (cf. section 4.3):

$$\rho_{12} = \frac{\phi_{y_1 y_2}}{\left[\phi_{y_1}\phi_{y_2}\right]^{\frac{1}{2}}}$$

and where

$$\phi_{y_1 y_2} = \frac{1}{T}\int_0^T y_1(t)\, y_2(t)\, dt \tag{4.32}$$

and

$$\phi_{y_i} = \frac{1}{T}\int_0^T y_i^2(t)\, dt, \quad i = 1,2.$$

Evidently, if $y_1(t) \equiv y_2(t)$, for all t in the range $(0 \leq t \leq T)$, then their cross-correlation ρ_{12} is one. The smaller this normalized cross-correlation, the more dissimilar the waveforms are, until, when $\rho_{12} = -1$, $y_2(t) = -y_2(t)$. Thus, one measure of the similarity between two waveforms is their cross-correlation coefficient. Since, as we have already observed, the optimum receiver bases its decision as to the waveform received on this measure, it is clearly the quantity of concern here.

So we have restated the problem as follows: We wish to find a mapping from the N possible signal amplitudes to N waveforms such that the cross-correlation between any two of these waveforms is as small as possible. The next section discusses some of the waveforms which may profitably be used for this purpose. Before proceeding, however, it is useful to obtain a standard whereby any N-level pulse communication system may be judged.

Presumably it is the maximum cross-correlation coefficient which will do the most damage since this is the measure of the similarity of the two waveforms most likely to be mistaken for each other. Accordingly, we might ask, what is the minimum value this maximum normalized cross-correlation coefficient can obtain? Suppose we have N waveforms: $x_1(t)$, $x_2(t)$, . . ., $x_N(t)$. Then the cross-correlation between any two of them is

$$\rho_{ij} = \frac{\int_0^T x_i(t)\, x_j(t)dt}{\left[\int_0^T x_i^2(t)dt\right]^{\frac{1}{2}}\left[\int_0^T x_j^2(t)dt\right]^{\frac{1}{2}}} \quad . \tag{4.33}$$

The average cross-correlation is clearly

$$\rho_{ave} = \frac{1}{N(N-1)} \sum_{\substack{i=1 \\ i\neq j}} \sum_{j=1} \rho_{ij} \tag{4.34}$$

since there are N(N-1) cross correlation coefficients ρ_{ij}, $i \neq j$. The ρ_{ii} terms are all equal to one and are excluded from the average because we are interested in the similarity between two different waveforms, not the trivial question of how similar a waveform is to itself. But

$$\rho_{max} = \max_{\substack{i\ j \\ i\neq j}} \rho_{ij} \geq \rho_{ave} = \frac{1}{N(N-1)} \left[\sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{ij} - N\right] \tag{4.35}$$

where the condition that $i \neq j$ has been replaced by subtracting the total contribution of the ρ_{ii} terms from the double summation. Substituting from equation 33 and interchanging the order of integration and summation, we find that

$$\frac{1}{N(N-1)}\left[\sum_{i=1}^{N}\sum_{j=1}^{N}\rho_{ij}-N\right]$$

$$=\frac{1}{N(N-1)}\left[\int_0^T\sum_{i=1}^{N}\left\{\frac{x_i(t)}{\left[\int_0^T x_i^2(t)dt\right]^{\frac{1}{2}}}\right\}\sum_{j=1}^{N}\left\{\frac{x_j(t)}{\left[\int_0^T x_j^2(t)\,dt\right]^{\frac{1}{2}}}\right\}dt-N\right] \tag{4.36}$$

The integral can be written

$$\int_0^T\left[\sum_{i=1}^{N}\left\{\frac{x_i(t)}{\left[\int_0^T x_i(t)\,dt\right]^{\frac{1}{2}}}\right\}\right]^2 dt \tag{4.37}$$

and since the integrand is never negative the integral is at least zero. Then

$$\rho_{max}\geq\rho_{ave}\geq-\frac{1}{N-1} \tag{4.38}$$

We shall use this relationship as a standard to which to compare of the set of waveforms to be discussed in the succeeding sections.

4.8. Orthogonal and Bi-orthogonal Waveforms

In this section we consider several methods for realizing a set of signals with satisfactorily small cross-correlation coefficients and investigate the performance obtainable with them. When N is large, Equation (38) states that the maximum cross-correlation between any two different waveforms in any set of N waveforms will be at least (approximately) zero. We will therefore demonstrate two simply generated sets of waveforms for which $\rho_{ij}\equiv 0$ for all $i, j = 1, 2, \ldots, N$, $i \neq j$. Such a set of waveforms is called an orthogonal set.

The first of these sets is the following:

$$f_i(t) = \sqrt{2} \sin (\omega_c + i \frac{k_1\pi}{T})t \quad \nu T \leq t \leq (\nu+1)T \quad i = 1,2,\ldots,N \tag{4.39}$$

where $\omega_c = \frac{k_2\pi}{T}$ and where k_1 and k_2 are non-zero integers. By definition

$$\rho_{ij} = \frac{2}{T} \int_{\nu T}^{(\nu+1)T} \sin (\omega_c + \frac{ik_1\pi}{T})t \sin(\omega_c + \frac{jk_2\pi}{T})t \, dt$$

$$= \frac{1}{T} \int_{\nu T}^{(\nu+1)T} \cos ((i-j)\frac{k_1\pi}{T})t \, dt - \frac{1}{T}\int_{\nu T}^{(\nu+1)T} \cos(2\omega_c+(i+j) \frac{k_1\pi}{T}) \, tdt$$

$$= \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \tag{4.40}$$

and the waveforms are indeed orthogonal. The modulation scheme using this waveform is commonly denoted orthogonal frequency-shift keying (orthogonal FSK). Both this set, and the one to be discussed next, can be generalized in a straightforward manner to include non-orthogonal waveforms and, in fact, are not even restricted to be finite in number. The most interesting situation, however, is the one we are now considering.

The second set of orthogonal waveforms is equally simple. It consists of N pulsed sinusoids of duration T/N seconds, the information being conveyed by the specific interval of time for which the pulse is not zero. The envelopes of the pulses are shown in figure 4.9 for the case in which N = 4.

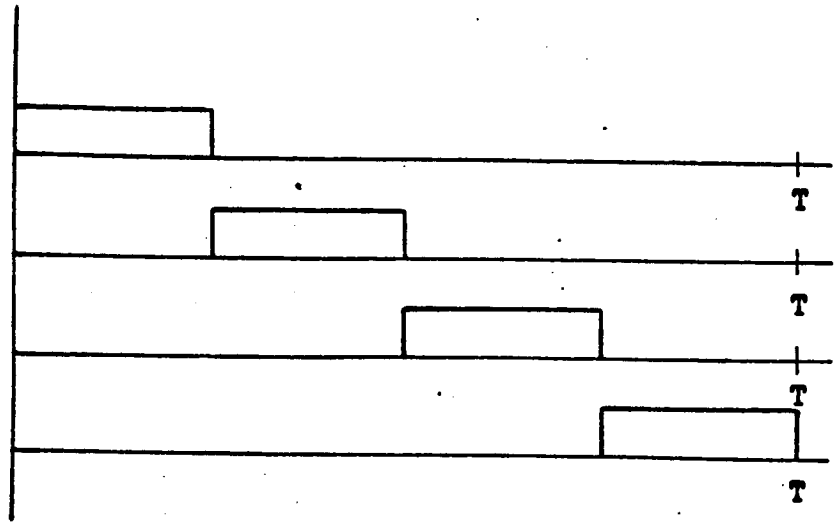


Figure 4.9. PPM Waveforms

Since no two different waveforms are non-zero at the same instant of time the cross-correlation coefficient ρ_{ij} is identically zero for all $i \neq j$ (see Equation 4.33). The waveforms can be denoted, for any N, by the equation

$$f_i(t) = \begin{cases} \sqrt{2} \sin \omega_c t & \frac{iT}{N} < t < (i+1)\frac{T}{N} \\ 0 & \text{otherwise} \end{cases} \qquad (4.41)$$

A communication device using these waveforms is said to be a <u>pulse-position modulation</u> (PPM) system.

The orthogonal encoder then accepts the data in blocks of n bits and, depending upon those bits, generates one of $N = 2^n$ waveforms to be transmitted. (In the examples given, each of the waveforms contained equal average power. This property will be assumed in the remainder of this discussion). If the waveform is correctly identified at the receiver, all n bits are transmitted without error. If a mistake is made in identifying the transmitted waveform, it is not difficult to show that, on the average, about one half of the n bits will be in error. To evaluate the performance of this system it is necessary, of

course, to determine the probability that an error will be made in identifying the received waveform. While this analysis is somewhat beyond the scope of these notes, the results have been sumarized in figure 4.10. The probability of an error in identifying a waveform at the receiver is plotted as a function of the "signal-to-noise ratio" $\frac{A^2T_B}{N_0}$ for various values of n, the number of bits represented by each waveform. The parameter $\frac{A^2T_B}{N_0}$, it will be recalled was the critical parameter which determined the threshold in conventional PCM. Here, as there, A^2 represents the average power in each waveform, while T_B is the time spent in transmitting each bit (the time per waveform is, therefore, nT_B) and N_0 is the noise spectral density. Note that the case n = 1 corresponds to uncoded PCM in which the normalized correlation between the two waveforms f(t) and g(t) is zero. As already stated, the probability of a bit error at the receiver is approximately one-half the probability that one waveform is mistaken for another.

It is seen that uniformly as n is increased the error probability is decreased for a given signal-to-noise ratio. In view of our experiences in analyzing other communication methods, we should expect to have to pay for this gain somehow, and when we turn our attention to bandwidth, we immediately discover where the toll has been extracted. Consider the case of PSK. We observed that the sinusoids had to be displaced from each other in frequency by an amount $k_1/2T$ where k_1 is integer greater than zero. If N such frequencies are used, N of these frequency slots are occupied, so that the effective bandwidth for orthogonal modulation is, letting $k_1 = 1$

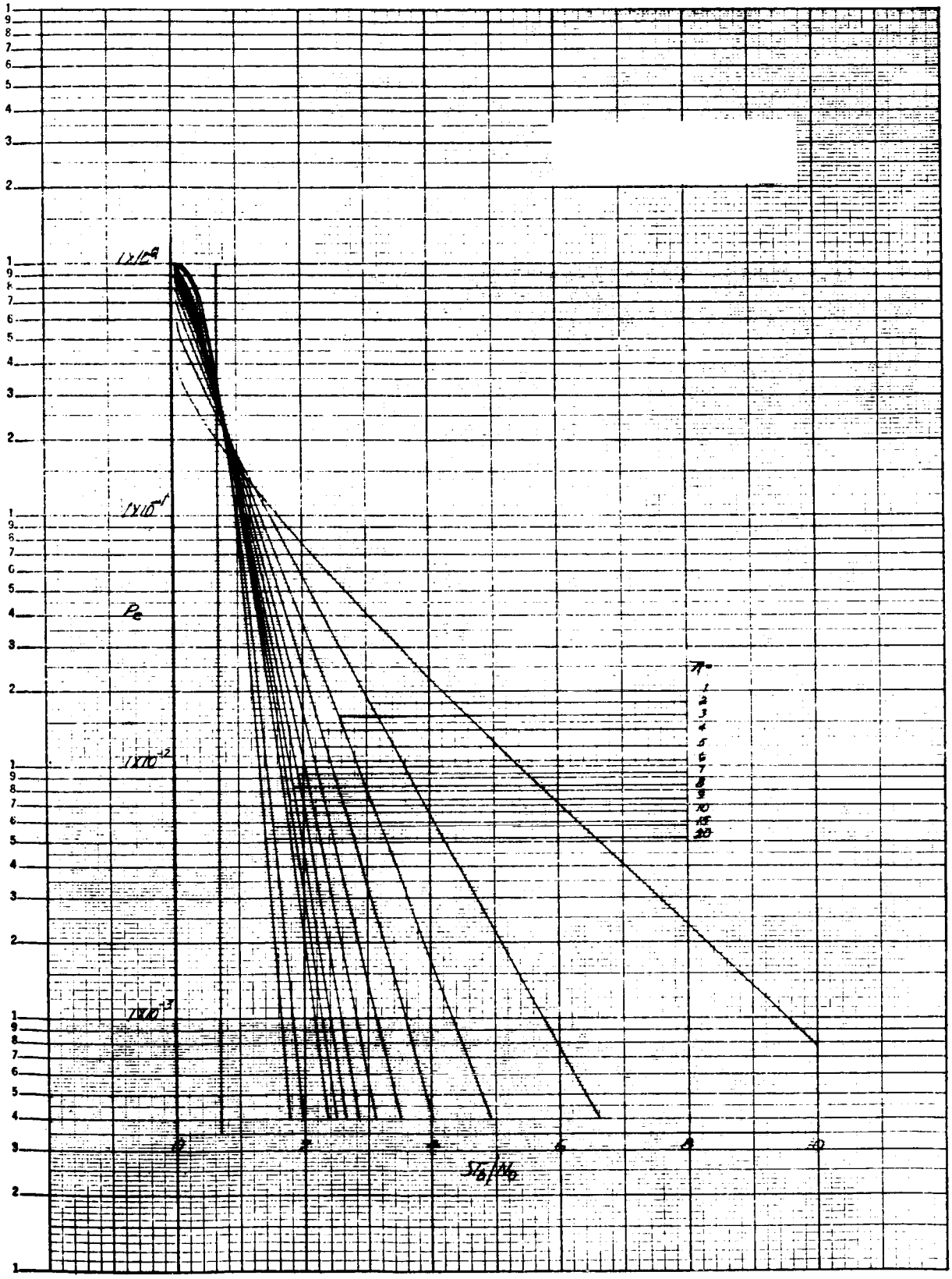


Figure 4.10. Error Probabilities for Orthogonal Codes

$$\left(B_{eff}\right)_{orthog.} = \frac{N}{2T} \tag{4.42}$$

Since n bits are transmitted in time T, $T = nT_B$, and since $(B_{eff})_{PCM} = \frac{1}{2T_B}$, the bandwidth expansion γ of orthogonal PSK relative to PCM is

$$\gamma = \frac{N}{n} = \frac{N}{\log_2 N} \tag{4.43}$$

Similarly, in the case of PPM, the pulses transmitted have a time duration of T/N where $T = nT_B$. Thus, the pulse width has been decreased relative to PCM by a factor n/N, and, as suggested in Section 1.5, the bandwidth must also be increased by the same factor N/n. This relationship holds in general for all orthogonal modulation systems.

A simple modification of orthogonal waveform sets is possible whereby the inequality on the average value of the cross-correlation

$$\rho_{ave} \geqslant -\frac{1}{N-1}$$

can actually be made an equality, while, at the same time, the bandwidth requirement is halved. The modification is to include not only the signals $f_i(t)$ but also $-f_i(t)$ in the set of permissible waveforms. This set of waveforms, called a <u>biorthogonal</u> set, has the property that

$$\rho_{ij} = \begin{cases} 1 & f_i(t) = f_j(t) \\ 0 & f_i(t) \neq \pm f_j(t) \\ -1 & f_i(t) = -f_j(t) \end{cases}$$

and it is easily verified that ρ_{ave} is indeed equal to $-\frac{1}{N-1}$ where N is the number of waveforms in the set. Moreover,

$$\left(B_{eff}\right)_{biorthog} = \frac{\frac{N}{2}}{2T} \tag{4.44}$$

since only N/2 frequencies, in the case of PSK, are needed to generate N biorthogonal waveforms. Similarly, if biorthogonal PPM waveforms are used, the pulse width can be T/(N/2). In both cases the bandwidth is only half that needed for orthogonal codes. Finally, since a waveform is even less likely to be mistaken for one with which its cross-correlation is minus one than it is for one to which it is orthogonal, the error probability is somewhat less, although the difference in this regard is rather insignificant. The probabilities of mistaking one waveform for another at the receiver, are plotted in figure 4.11, for biorthogonal modulation.

When n = 1, the biorthogonal set contains two waveforms, f(t) and -f(t). Evidently this is just the ordinary PCM system discussed earlier. Note that, if $\frac{A^2 T_B}{N_0}$ is greater than five, the probability of making an error in identifying the transmitted signal at the receiver is less than .001. It is this consideration that was used to arrive at the requirement of Equation (4.24). When n = 6, it is seen that the same performance can be attained with a signal-to-noise ratio as small as three. This is accomplished using biorthogonal$_6$ codes at the expense of an increase in bandwidth by a factor of $2/12 = 5\frac{1}{3}$ over that needed for ordinary PCM. The effective gain of $\frac{5}{3}$ in power may well be worth the price.

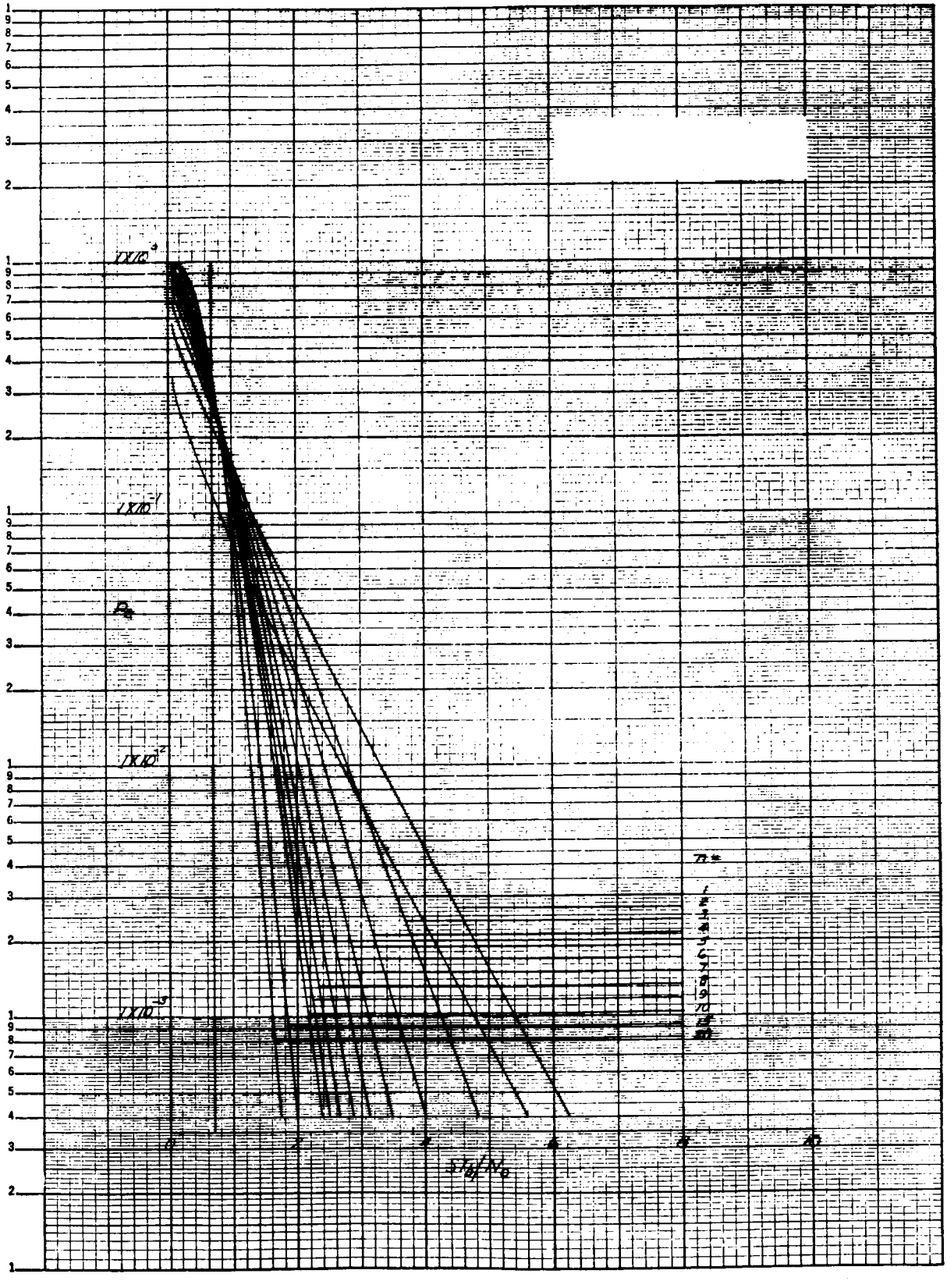


Figure 4.11. Error Probabilities for Bi-Orthogonal Codes

4.9 Synchronization

It will be noted that all of the systems analyzed have depended upon a common time reference (i.e. synchronization) for the transmitter and receiver. In particular, it was assumed that it is possible to generate a sinusoid at the receiver which has the exact phase and frequency of the received carrier. This information is usually called carrier synchronization. Further, optimum demodulation depends on the knowledge of the instants of time in which a waveform begins and ends (i.e. the instants of time νT, $\nu = 0, 1, \ldots$.) This is called word or waveform synchronization. Other timing references are often necessary, too. If, for example, the waveform conveys n bits of information where each sample is quantized to $M=2^m$ levels, as in the previous section, the data user must have available synchronization information which will enable him to separate the received sequence of bits into the right blocks of m bits. In addition, if, as is almost always the case, numerous data sources are to be multiplexed together, the user must be able to identify which part of the data corresponds to which source. This is referred to as frame synchronization.

Numerous synchronization techniques have been explored. One of the most common and straightforward methods involves the use of a phase-locked loop to track an unmodulated sinusoid, or other periodic signal, the phase and frequency of which convey the necessary timing. We shall delay further discussion of the synchronization problems until Section Six in which we investigate the related problem of ranging.

5 DATA COMPRESSION

All of our efforts in investigating methods for improving the efficiency of long distance communication have been concerned with the communication system itself. But there is another area in which great improvements can theoretically be made, viz. in the pre-processing of the data which are to be communicated. The sampling theorem states that the data must be sampled at least every $T = \frac{1}{2W}$ seconds. Unfortunately, W is not generally known for the data sources being monitored. Indeed, it is precisely the uncertainty and the difficulty in predicting the behavior of the sources that make them worthwhile to observe. Thus, W, and therefore the sample rate T, must be estimated, and in order to insure that important information is not lost, they must be conservatively estimated with the result that a source is usually sampled considerably more rapidly than it would ideally be. In addition, even if W were known and the sampling rate optimally determined, successive samples would, in general, be rather highly dependent. That is, much of the information contained in the $(i+1)^{st}$ sample would have been predictable from the i^{th}, $(i-1)^{st}$ $(i-2)^{nd}$, etc. samples. If these samples themselves are transmitted, much effort is wasted in sending <u>redundant</u> information. The effects of redundancy is readily observed in written text. CNSDR FR XMPL THS SNTNC. Few will find any difficulty in reading that array of letters even though more than 30% of the symbols have been deleted. That much greater elimination of redundancy is possible is evidenced by the various shorthand and speed

writing techniques which have been developed. Any telemetry system that transmitted conventionally written text would be sending much redundant information and thereby inefficiently using its power.

The suggestion then is that shorthand techniques should be used aboard a spacecraft so only essential information is transmitted. Since fewer symbols would be sent, more time could be spent per symbol, thereby decreasing the probability of an error at the receiver and hence increasing the reliability of the system.

This argument for the removal of redundancy seems at variance with the discussion of the previous chapter in which it was stated that, at least under certain conditions, redundant symbols could be advantageously added to reduce the likelihood of an error at the receiver. The difference in the two situations is in the fact that the added redundancy can be carefully selected for the maximum improvement whereas the naturally occurring redundancy, while providing some error-correcting ability, is far from efficient. Suppose, for example that the following sequence of temperatures were received: 250.031° 250.033° 250.035° 350.033° 250°.034° ..... . The error in the most significant position of the fourth reading is easily corrected. Yet an error in the least significant position would probably pass unnoticed. As an alternative method suppose only the first reading were sent followed by the incremental temperature changes, each repeated three times as follows: 250.031°, +2 +2 +2, +2 +2 +2, -2 -2 +2, +1 +2 +1, ... The same number of symbols are involved in both cases(actually, the sign is one of only two symbols, rather than one of ten), yet the error in the least significant position of the fourth and fifth readings is readily discernable. In the latter case, the natural redundancy was replaced with controlled

redundancy, thereby increasing the correctability of small errors. This system is not necessarily practical. Changes of more than 9° cannot be indicated, and in addition, if an error does occur, all succeeding readings will be in error since the incremental changes are transmitted rather than the readings themselves. Nevertheless, it hopefully does illustrate the point. This chapter will discuss several techniques for the elimination of redundancy, or as it is variously called, data compression or data compaction.

5.1 N^{th} Differences

In the example of the preceding section, the incremental changes or first differences of the data samples were transmitted rather than the samples themselves. This method can be extended by transmitting the second differences (i.e., the first differences of the first differences) or third differences, etc. The advantage of this procedure, of course, depends upon the data. If, for example, the data are increasing, or decreasing, nearly linearly, $x_\nu \approx a + b\nu$, then the first differences are nearly constant, the second differences are nearly all zero. Transmission of the second differences in this case would generally require no more than one symbol. In general, if the data are varying approximately as an $(n-1)^{st}$ polynomial in time, $x_\nu = a_1 + a_2 \nu + \ldots + a_n \nu^{n-1}$, then the n^{th} differences are nearly constant and and can be transmitted with a very few symbols. It will be observed that, in order to reconstruct the original data from the n^{th} differences, the first n data samples (or their equivalent) must also be transmitted. For example, if third differences are transmitted, it is seen, referring to figure 5.1, that all the data samples can be reconstructed if either the underlined or the bracketed

information is transmitted.

Samples	First Differences	Second Differences	Third Differences
$[x_1]$	$[x_1 - x_2]$	$[x_1 - 2x_2 + x_3]$	$[x_1 - 3x_2 + 3x_3 - x_4]$
x_2	$x_2 - x_3$	$x_2 - 2x_3 + x_4$	$[x_2 - 3x_3 + 3x_4 - x_5]$
x_3	$x_3 - x_4$	$x_3 - 2x_4 + x_5$	$[x_3 - 3x_4 + 3x_5 - x_6]$
x_4	$x_4 - x_5$	$x_4 - 2x_5 + x_6$	
x_5	$x_5 - x_6$		
x_6			

Figure 5.1 A third difference data compression scheme

The disadvantages of the n^{th} differences scheme are apparent. First of all, because of the tendency of errors to propagate, as observed earlier, the process must be periodically truncated and begun again. But more seriously, it is seldom possible to predict that the data will closely approximate an n^{th} order polynomial for any particular value of n. Because the data does frequently have a constant non-zero mean, and because it must generally be sampled too often in order to insure that no information is lost, the first difference schemes are sometimes practical, but seldom are higher-order difference schemes used.

5.2 Run Length Encoding

It has been observed that, because of the lack of knowledge concerning the bandwidth of the data sources, the data must generally be sampled too

often. The redundancy can be considerably reduced by taking the first differences of the samples. Even so, however, because the data bandwidth is difficult to predict, it is not known in advance how different successive samples will be. Thus, it is not known how many symbols should be allowed for each first difference; i.e., will two successive samples most likely differ only in the least significant position or the two least significant positions, or will they sometimes differ in all positions? Presumably, all of these situations will occur at various times. But if the first differences demand nearly as many symbols as did the original samples, there is little advantage in taking first differences.

One method for partially overcoming this difficulty is by the technique of run-length encoding. Suppose the data samples are quantized and represented by binary digits. The first differences are also binary digits, but since presumably two successive samples do not usually differ significantly, the first differences will consist mainly of zeros, particularly in the most significant positions. The run-length encoding scheme takes advantage of this fact by transmitting, not the first differences themselves, but rather the spacings between successive "ones" in the binary sequence representing these first differences. Thus suppose the data samples are 0 1 0 1 1 1 0 1, 0 1 0 1 1 1 0 0, 0 1 0 1 1 0 1 1, 0 1 0 1 1 0 1 1, 0 1 0 1 1 1 0 0 ..... where the first digit represents the sign of the sample. The first differences are then

..... 1, 0 0 0 0 0 0 0 1, 0 0 0 0 0 0 0 1, 0 0 0 0 0 0 0 0 , 1 0 0 0 0 0 1, ... and the run-length code, obtained by counting the spacing between successive ones in the first difference sequences and representing the spacings in binary form, is 1 1 1, 1 1 1, 1 0 0 0, 1 1 0, ...

The number of digits has thereby been reduced from 32 to 13 without sacrificing the ability to transmit an eight digit first difference if necessary. This example is somewhat misleading since we have ignored the necessity of providing the commas in the run-length encoded sequence. One not particularly efficient method for overcoming this difficulty is to keep the encoded sub-sequence length constant. That is, suppose we divide the code into groups of three digits and constrain it so that no run of zeros greater than six can be represented by one set of three digits. If the binary number seven is transmitted, the sequence immediately following it is to be added to it to determine the distance between successive ones in the original sequence. Thus, if the distance between successive ones is six the sequence 1 1 0 is transmitted; seven is represented by 1 1 1, 0 0 0; eight by 1 1 1, 0 0 1; thirteen by 1 1 1,1 1 0; fourteen by 1 1 1, 1 1 1, 0 0 0, etc. The data sequence example used previously is now encoded as 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 1 1 1 0 now requiring 21 symbols. No special information is needed here to decode this sequence however, other than the knowledge of its beginning, since it is always divided into groups of three. It will be noted that this scheme is not constrained to be used only in conjunction with the first differencing technique. It is applicable whenever long runs of zeros (or ones) occur frequently in the information to be transmitted.

5.3 Huffman Codes

Another method for reducing the average numbers of digits to be transmitted without sacrificing the ability to transmit all possible messages lies in the use of Huffman codes. Observe in the example of the last section that the most commonly occurring first difference sequences are apparently 0 0 0 0 0 0 0 0, 0 0 0 0 0 0 0 1 and 1 0 0 0 0 0 0 1. In probable des-

cending order of occurrences the remaining sequences are:

0 0 0 0 0 0 0 1 0, 1 0 0 0 0 0 0 1 0, 0 0 0 0 0 0 1 1, 1 0 0 0 0 1 1,

0 0 0 0 0 1 0 0, 1 0 0 0 0 1 0 0, etc. The rationale for the Huffman encoding process is as follows: The average number of symbols necessary to represent the data will be reduced by representing the most common sequences by fewer symbols than the less common sequences. This is best illustrated by an example. Suppose the sequences

0 0
0 1
1 0
1 1

occur with the relative frequencies 1/2, 1/4, 1/8 and 1/8 respectively. If the sequences themselves are transmitted the average number of digits per sequence is, of course, two. But suppose the following identification is made

0 0 → 1
0 1 → 0 1
1 0 → 0 0 1
1 1 → 0 0 0

Then the average number of symbols transmitted is (with $P_r(x)$ indicating the probability or relative frequency of the event x)

$$1\ Pr(0\ 0) + 2\ Pr(0\ 1) + 3\ Pr(1\ 0) + 3\ Pr(1\ 1)$$

$$= 1.75,$$

not a vast improvement perhaps, but one which becomes considerably more impressive as the sequence length increases and the difference between the probability of the most probable sequence and that of the least probable increases. It may be wondered why the particular identification of the expression (1) was chosen rather than some other. Why, in particular, weren't more

of the two symbol sequences used? The reason is because any random ordering of the symbol sequences on the right side of the expression (1) can be uniquely deciphered. This follows from the observation that each symbol either ends in a "one" or contains three zeros. Three consecutive zeros are therefore recognized as corresponding to the sequence 1 1, while the end of any other symbol will be identified by a one. This is not true for all possible mappings of the form (1). The Huffman coding algorithm, however, guarantees that such a mapping can always be accomplished. The optimum mapping depends upon the probabilities of occurrence of the original sequences. Nevertheless, even if these probabilities are not known, some saving can generally be effected by mapping those first-difference sequences corresponding to the least change between successive samples onto the shortest symbol sequences, those corresponding to greater differences onto longer sequences. This method, like the one of the previous section, clearly is not restricted to be used only on the first differences of the data. It is useful whenever the different sequences which are to be transmitted occur with different probabilities.

5.4 The Floating Barrier Technique

Another method for data compression which we shall consider briefly in this far from exhaustive discussion is called the floating barrier scheme. With this method a quantization level is set and an initial sample is transmitted. The data are periodically sampled, as before, but successive samples are transmitted only if they differ from

the last transmitted sample by more than the pre-determined quantization. The operation of the process is illustrated in figure 5.2. The sampling interval is Δt and the quantization amplitude is q. The transmitted samples are encircled

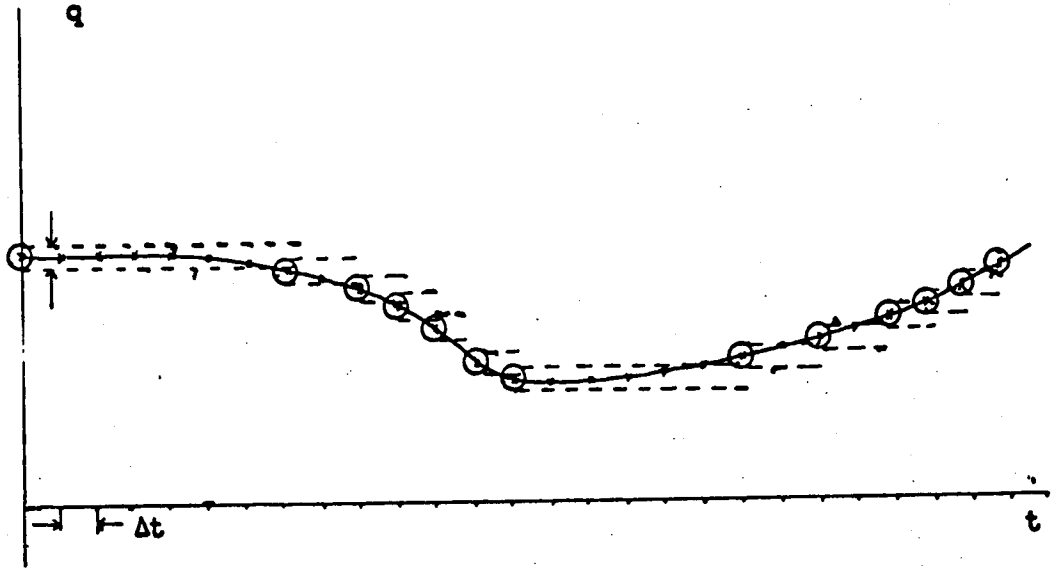


Figure 5.2 The floating barrier scheme

It is seen that for this particular example the number of samples which must be transmitted is reduced from 27 to 13, yet the signal is always known to be within $\pm$ q/2 of the sample points.

It is not sufficient here just to send the samples, however. Since every sample is not transmitted it is also necessary to provide the number of samples which were omitted between successive transmitted samples if it is to be possible to reconstruct the original signal. This, of course, somewhat decreases the savings in the number of digits which must be transmitted.

5.5 Television Compression

Before leaving the subject of data compression it is well to consider one case in which some rather impressive savings can be made, namely, in the transmission of video information. It has been verified that the intensity of the light representing a television picture can be quantized to only eight levels from black to white with little picture degradation. If there are sixteen quantization levels the eye can see the difference between the quantized picture and one in which a continuum of intensity levels is allowed only with difficulty. A commercial television picture is quantized spatially into 512 horizontal lines. If it were further quantized vertically, also into 512 lines for example, a picture could be represented in digital form as an array of 512 x 512 numbers which, assuming eight-level intensity quantization, could vary from 0 to 7, or in binary form, from 0 0 0 to 1 1 1. Thus it would take 512 x 512 x 3 = 786, 432 binary digits to represent one television picture. But, clearly, these digits are not independent. A black spot is likely to be surrounded by black spots, a white spot by white spots. Furthermore, on the relatively rare occasions in which two adjacent spots do differ they most generally differ by only one intensity level. This dependence exists, of course, not only spatially but temporally if successive pictures of the same object are transmitted, and not only between adjacent symbols but over a considerably greater distance. Clearly, then there is room for a vast reduction in the amount of data which actually must be transmitted. By what factor

the data may be compressed and by what techniques it may be accomplished depends to a large degree on the assumptions made about what is being observed. Reductions by a factor of five to ten in the number of bits necessary to represent a picture seem to be well within the realm of practicality. Theoretically even more impressive reductions are possible.

5.6 Destructive Compaction

The methods of data compression we have considered up to now are sometimes referred to as non-destructive data compression techniques. All the information from the source is to be transmitted; only redundancy is to be removed. Often, however, the user is not interested in all the available information. Perhaps it is only the maximum and minimum temperatures of some device for example, that are of interest. If this is the case, it is certainly wasteful to send the complete set of temperature readings. It is relatively easy to determine the extremes on board the spacecraft and to send those readings only. Similarly, if it is desired to know only the statistical distribution of the counts from a Geiger counter it is certainly not efficient to transmit a complete record of the number of counts per second. Rather, a data reducing mechanism aboard the vehicle should be used to process the information to determine the desired histogram which can then be transmitted.

This sort of data reduction is called destructive data compression since not all of the information provided by the source is used. The percentage savings can be extremely high for this kind of data compression in cases in which the user needs only a quite specific subset of the information which could be provided. As an extreme case, suppose it were

necessary to know the average geiger counter reading over a period of 10 hours. The amount of transmitted data could be reduced by a factor of 36,000 by sending only this average rather than sending, for example, a reading every second.

The limitation of this procedure, however, lies in the fact that the user seldom knows what aspect of the data will be useful to him before he sees it. He is generally reluctant to allow any information to be destroyed for fear of missing some important but unsuspected observation.

5.7 Summary and Conclusions

Numerous methods have been investigated for the elimination of redundancy in data which is to be telemetered. Among them are included:

(1) The n^{th} difference method whereby the data samples themselves are not transmitted but rather their first, second, or higher order differences.

(2) The process of run-length encoding in which the presumed predominance of zeros in the data (or in their first or higher order differences) is used to reduce the number of symbols transmitted.

(3) Huffman encoding which relies upon the generally valid assumption that some data samples (or commonly, data sample first-differences) occur more frequently than others.

(4) The floating barrier technique which transmits samples only when they are significantly different from the last transmitted sample.

(5) Destructive compression techniques.

The amount of reduction in the transmitted data afforded by compaction varies widely from source to source. Compaction ratios of between two and ten are fairly common with non-destructive methods; destructive compression ratios are almost unlimited depending upon the amount of information actually needed.

Each of the methods discussed involves several difficulties. One of the more serious is the tendency of errors to propagate. In order that (non-destructive) data compression be significant a considerable amount of statistical information concerning the source* must be available (information which is generally not entirely available when applied to spacecraft sources since each experiment encounters rather unpredictable situations). This coupled with the fact that data compaction equipment considerably complicates the spacecraft electronics,** has in the past limited data compaction efforts. Nevertheless, as data handling capabilities of spacecraft are to be significantly increased, data compression techniques will become more and more attractive.

---

*This fact has given considerable impetus to the investigation of self-adaptive data compression schemes which have the ability to vary in accordance with the source statistics which are estimated on board the spacecraft.

**All of the above techniques, for example, are more effective when the data are changing slowly than when they are rapidly changing. Thus the data rates will vary depending upon the sources. Since the transmitter must generally operate at a fixed rate, buffering or temporary storage facilities must be provided.

6 SPACECRAFT TRACKING

The problem of tracking or keeping a running record of the position and velocity of a spacecraft is not generally classified as a telemetry problem. The word telemetry usually implies the transmission of information through space from one position to another.

Yet, spacecraft tracking involves the same equipment and many of the same techniques as does telemetry; it logically falls under the study of telemetry systems.

Tracking can be accomplished with one antenna or several. If two or more antennas are involved, it is theoretically possible to obtain all information concerning the position of the vehicle by pointing all the antennas at it and observing their respective orientations. The antennas may "see" the vehicle either from its radio transmission or by its reflection of radar signals from the ground. This technique suffers from two disadvantages. First, it does require two or more antennas suitably placed, with a means of communicating to a common point the orientations of each antenna, in order to determine the position of the object being observed. In addition, it is difficult to obtain the desired accuracy when the spacecraft is fairly distant from the earth. Even a 0.01° degree error in the orientation of each of the antennas can result in a sizable error in the estimate of the position of the spacecraft as shown in figure 6.1. Since a successful mission strongly depends upon the ability to place the spacecraft on

the desired trajectory, which in turn depends upon the precision with which its position can be measured, more precise methods must be considered.

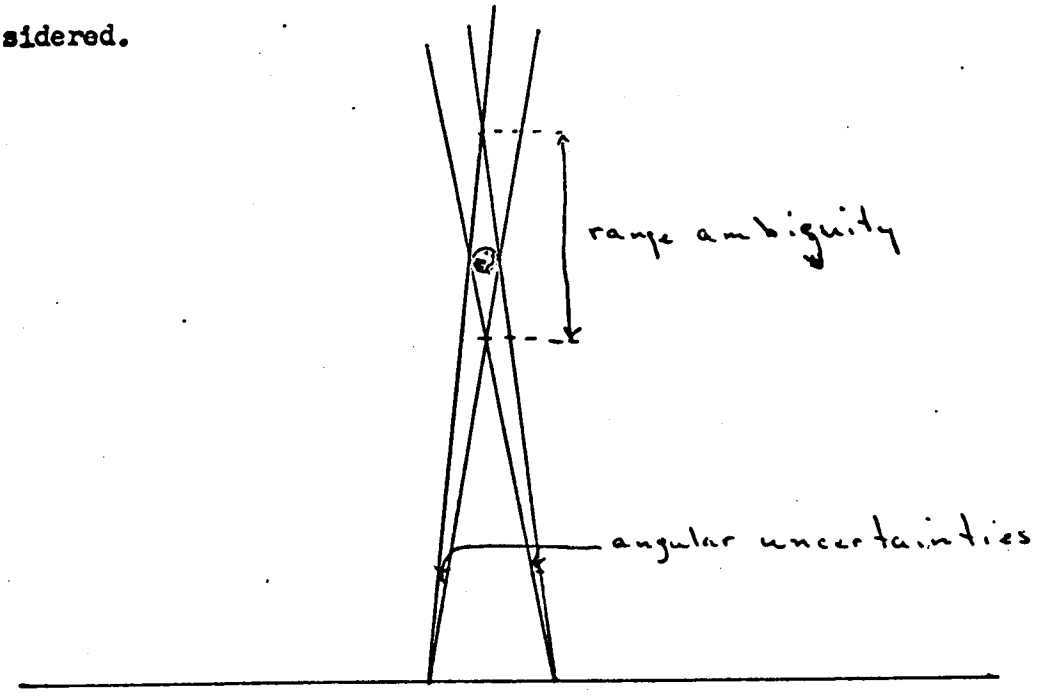


Figure 6.1 Range Ambiguity in Triangulation

Since, as seen in figure 6.1, the distance from the earth, in particular, is subject to large ambiguities when it is determined from triangulation , it is apparent that another means of determining this distance or range would be useful. Such a technique, long used in conjunction with conventional radar is to transmit a narrow pulse towards the vehicle and measure the time which elapses before the echo returns. Since these time delays can be measured electronically to a very high precision, the distance can be determined quite accurately. The ultimate limitation in measuring distance by this method lies in the precision with which the velocity of electromagnetic radiation is known.

Several practical improvements can be made on this technique of ranging. In order to understand these improvements it is necessary to consider

more precisely how ranging information is extracted from a reflected pulse train.

6.1 Pulsed Radar

In Chapter Four we were introduced to the concept of correlation detectors. It was argued that the optimum detector for estimating which of the signals $x_i(t)$ $(i = 1, 2, \ldots)$ was most likely transmitted when the received signal was $y(t)$ involved the formulation of the integrals

$$I_i = \int y(t)\, x_i(t)\, dt \tag{6.1}$$

and the selection of the largest of these.

The determination of range involves essentially the same considerations. Here, assuming the transmitted pulse is rectangular

$$x(t) = \begin{cases} \sqrt{2}\, B \sin \omega_c t & 0 < t < \Delta t \\ 0 & \text{otherwise} \end{cases} \tag{6.2}$$

the received signal is one of the signals

$$y_s(t) = \begin{cases} \sqrt{2}\, A \sin \omega_c (t - t_0) & t_0 < t < t_0 + \Delta t \\ 0 & \text{otherwise} \end{cases} \tag{6.3}$$

where t_0 can have any value over the range of ambiguity, $t_1 < t_0 < t_2$. The optimum detector determines the integrals

$$I(\tau) = \int_{\tau}^{\tau + \Delta t} y(t)\sqrt{2}\sin \omega_c (t - \tau)\, dt \qquad (6.4)$$

where $y(t) = y_s(t) + n(t)$, $n(t)$ representing the additive noise. The best estimate of the delay t_o is the value of τ maximizing the integral (4).

The contribution of the signal $y_s(t)$ to integrator output is illustrated in Figure 6.2 as a function of τ. Hopefully the integrator will attain its maximum value for $\tau x t_o$ even when noise is present. In order to estimate the signal power needed to provide reasonable assurance of this happening, consider the probability that the output $I(t_o + \Delta t)$ actually exceeds $I(t_o)$. Clearly, this probability must be quite small if the delay is to be determined at all accurately. It is not difficult to see that the probability of $I(t_o + \Delta t)$ exceeding $I(t_o)$ is identical to the probability of an error in the detection of one of two orthogonal signals of average power A^2 and time duration Δt. Thus, referring to Figure 4.10, we find that if $A^2\Delta t/N_o = 8$, the probability of $I(t_o + \Delta t)$ exceeding $I(t_o + \Delta t)$ is approximately 2.4×10^{-3}. This provides a lower band on the signal power required for accurate ranging.

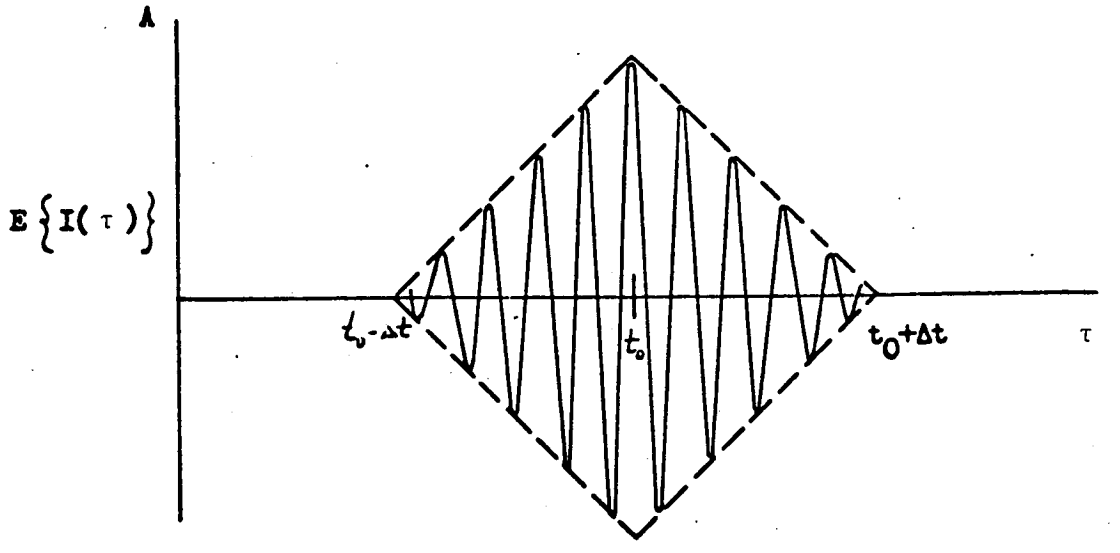


Figure 6.2. Pulse-Matched-Filter Output

The primary limitation to this procedure is in the amount of peak power the transmitter is able to radiate. Accurate ranging requires Δt to be small, while, at the same time, the product $A^2\Delta t/N_0$ must be large. This limitation can be partially overcome by sending the pulses periodically with a repetitive rate T which is greater than the range of ambiguity, $T > t_2 - t_1$. Then the summation

$$\sum_{i=0}^{N-1} I_i(t_0) = \sum_{i=0}^{N-1} \int_{\tau}^{\tau+\Delta t+iT} y(t) \sqrt{2} \sin \omega_c(t-\tau)\,dt \qquad (6.5)$$

can be used to estimate t_o, increasing the effective integration time, and, hence, the ratio $A^2\Delta t/N_0$ by a factor of N. In order to do this, the vehicle being ranged must either be relatively stationary or else the relative motion between the receiver and the vehicle during the interval between pulses must be known and compensated for. This latter can be accomplished by transmitting in addition to the pulse, some unmodulated sinusoid at the frequency ω_c. This sinusoid can be tracked, for example, with a phase-locked loop and the loop output used to control the clock generating the pulses as shown in Figure 6.3.

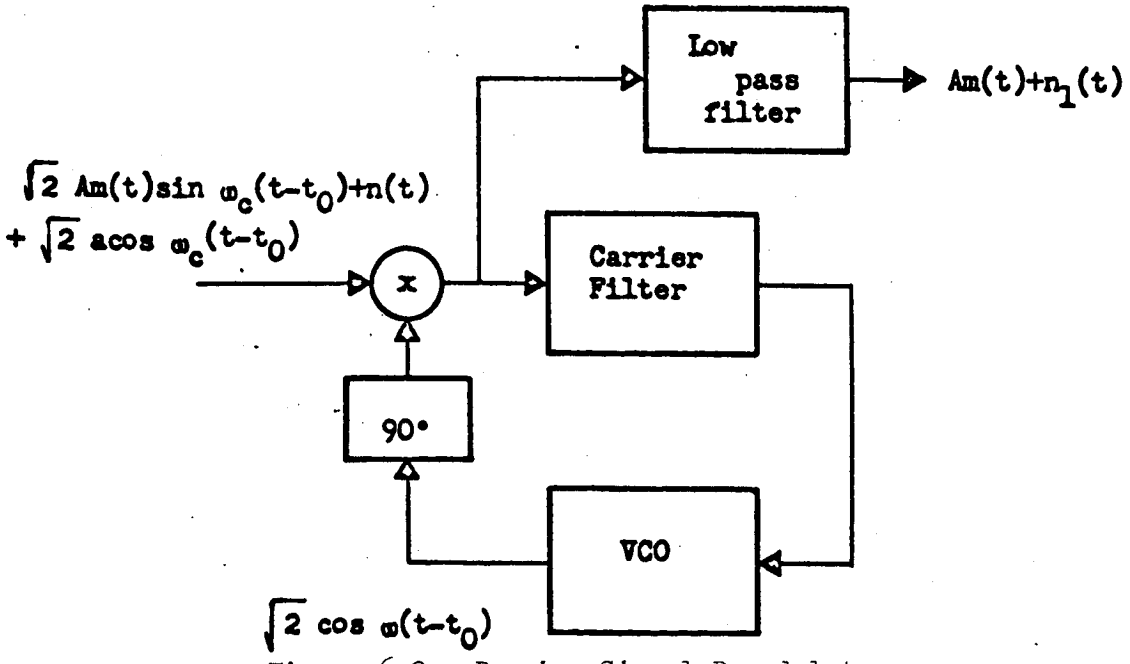


Figure 6.3. Ranging Signal Demodulator

If the tracked sinusoid is also used to demodulate the pulse before it is integrated, as shown in Figure 6.3, the integrals of Equation (6.4) become

$$I'(\tau) = \int_{\tau}^{\tau+\Delta t} y'(t)\, dt \tag{6.6}$$

where $y'(t) = Am(t) + n_1(t)$,

$$m(t) = \begin{cases} 1 & t_0 < t < t_0 < \Delta t \\ 0 & \text{otherwise} \end{cases}$$

The continuum of integrals can then be determined with one integrator followed by a sampler

$$\begin{aligned} I'(\tau) &= \int^{\tau+\Delta t} y'(t)\, dt - \int^{\tau} y'(t)\, dt \\ &= Y(\tau + \Delta t) - Y(\tau) \end{aligned} \tag{6.7}$$

where

$$Y(t') = \int^{t'} y'(t)\, dt \tag{6.8}$$

The locally generated signal $\cos \omega_c(t - t_0)$ can often be used to control the clock gating the sampler which samples the output of the integrator. If the carrier phase and the pulse phase are coherent, it is only necessary to sample at the instants of time $\hat{t}_0 = 2\pi i/\omega_c$, $i = 1, 2, \ldots$, since a pulse cannot begin at any other instant of time.

The primary disadvantage to this ranging procedure as observed is that all the signal power must be sent in a relatively short period of time. If the pulse repetition rate is T and the pulse width is Δt, pulse energy is being transmitted only $\Delta t/T \times 100\%$ of the time. The peak power radiated must be increased by a factor of $T/\Delta t$ if the average is to be kept the same as that

which would be possible with continuous radiation.

But how can continuous radiation be used to achieve the desired range resolution? One answer lies in the use of pseudo-random sequences.

6.2 Pseudo-Random Sequences

We have observed that the optimum receiver for determining range involves the formulation of the integrals

$$I(\tau) = \int_{\tau}^{\tau+N\Delta t} y(t)\, x\,(t - \tau)\, dt \tag{6.9}$$

for all contending values of t_0, where $y(t)$ is the received noise corrupted signal and $x(t - \tau)$ is the locally generated expected signal starting at the time τ and lasting for a duration $N\Delta t$. (The significance of the factor N will become apparent shortly. In the preceding section, of course, N = 1.) In the absence of noise

$$I(\tau)_{n(t)=0} = A\int_{\tau}^{\tau+N\Delta t} x(t - t_0)\, x\,(t - \tau)\, dt$$

$$= A\int_{0}^{N\Delta t} x(t + \eta)\, x(t)dt \quad A\phi(\eta) \tag{6.10}$$

where $\eta = \tau - t_0$. The function $\phi(\eta)$ is called the autocorrelation function of $x(t)$. Suppose $x(t)$ consists of a sequence of N pulses of width Δt as shown in Figure 6.4. Further suppose it is transmitted continuously so that, after the N^{th} pulse is transmitted, the sequence is immediately retransmitted beginning with the first pulse. It is easily verified that, in this case,

$$\phi(\tau) = \phi(n\Delta t)\left[\frac{(n+1)\Delta t - \tau}{\Delta t}\right] + \phi\left[(n+1)\Delta t\right]\left[\frac{\tau - n\Delta t}{\Delta t}\right] \quad n\Delta t \leq \tau \leq (n+1)\Delta t \tag{6.11}$$

The details of the proof of this statement (6.11) are left as an exercise for the reader.

This observation allows us to consider the autocorrelation function only at the points $n\Delta t$, $n = 0, 1, \ldots$, since we know that $\phi(\tau)$ varies linearly between these points. (Note that $\phi(\tau) = \phi(-\tau)$.) Moreover, a pulse train of the type considered here can be represented simply by specifying the amplitude of the successive pulses. Thus, for example, the pulse train in figure 6.4 can be represented by the sequence $a_1\ a_2\ a_3 \ldots a_N$

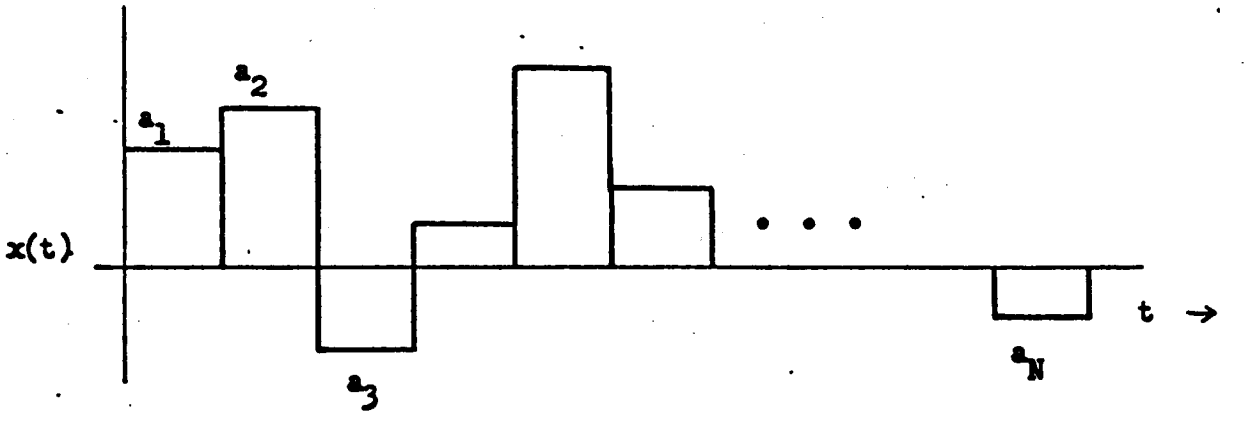


Figure 6.4. A Sequence of Fixed-Amplitude Pulses

where a_i denotes the amplitude of the i^{th} pulse. Since the pulse train is periodic with period $T = N\Delta t$, $a_{N+j} = a_j$. The autocorrelation function $\phi(i\Delta t)$ is therefore

$$\phi(i\Delta t) = \int_0^T x(t)\, x(t+i\Delta t)\, dt$$

$$= \Delta t \sum_{j=1}^{N} a_j\, a_{j+i} \qquad (6.12)$$

and it, too, can be readily determined from the knowledge of the amplitudes of the successive pulses.

Consider now the periodic pulse train represented by the sequence 1, 1, -1 as shown in figure 6.5(a). The autocorrelation function is easily determined

$$\begin{aligned}
\phi(0) &= \Delta t \sum_{j=1}^{3} a_j^2 = 3\Delta t \\
\phi(\Delta t) &= \Delta t \sum_{j=1}^{3} a_j a_{j+1} = -\Delta t \\
\phi(2\Delta t) &= \Delta t \sum_{j=1}^{3} a_j a_{j+2} = -\Delta t
\end{aligned} \tag{6.13}$$

and repeats with period $T = 3\Delta t$. The autocorrelation function is illustrated in figure 6.5(b).

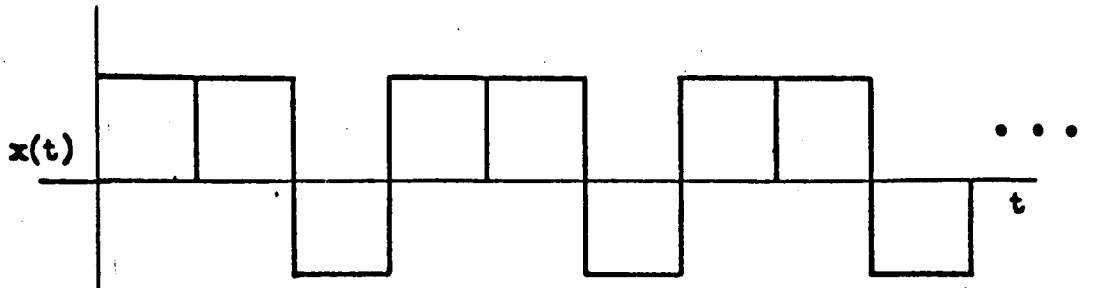


Figure 6.5(a) The periodic pulse train x(t)

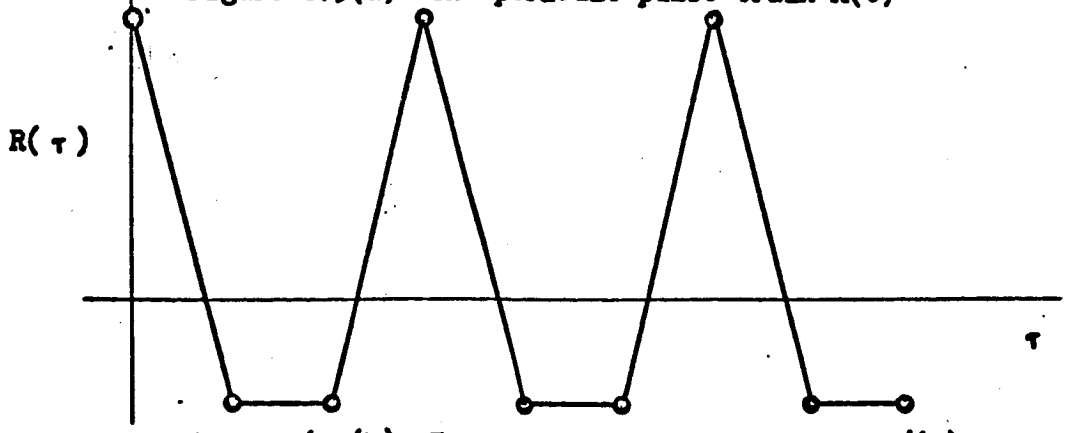


Figure 6.5(b) Its autocorrelation function $\phi(\tau)$

The reader may verify that the sequence

1 1 1 -1 -1 1 -1

when repeated periodically has the autocorrelation function

$$\phi(0) = 7\Delta t$$

$$\phi(i\Delta t) = -\Delta t \qquad i = 1, 2, 3, 4, 5, 6.$$

It is, in fact, possible to construct sequences whose periodic autocorrelation function has the values

$$\phi(0) = N\Delta t \tag{6.14}$$

$$\phi(i\Delta t) = -\Delta t \qquad i = 1, 2, \ldots, N-1$$

for all values of $N = 2^n - 1$, $n = 1, 2, \ldots$. These sequences are called pseudo-random or PR sequences because of the many properties which they have in common with truly random two-level sequences.

The reader may have observed that any one of these sequences, combined with all its permutations, comprise a set of N wave forms whose maximum normalized cross-correlation is $\frac{\phi(i\Delta t)}{\phi(0)} = -1/N$, $i \neq 0$, thereby nearly attaining the bound on the maximum cross-correlation derived in section 4.7. An additional property of these sequences is that they contain $\frac{N+1}{2}$ pulses with the amplitude +1 and $\frac{N-1}{2}$ pulses with the amplitude -1. Thus the cross-correlation between any phase shift of one of these sequences with the sequence consisting of only -1's is

$$-\Delta t \sum_{i=1}^{N} a_i = -\Delta t$$

Consequently the set of all permutations of such a sequence plus the sequence consisting of all -1's contains N + 1 sequences with

the maximum normalized cross-correlation equal to -1/N, and, hence, actually does attain the bound derived in section 4.7. These sequences can be used to generate waveforms in the manner described there. To illustrate, consider the set of sequences

$$\begin{matrix} -1 & -1 & -1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{matrix}$$

which represent the waveforms shown in figure 6.6.

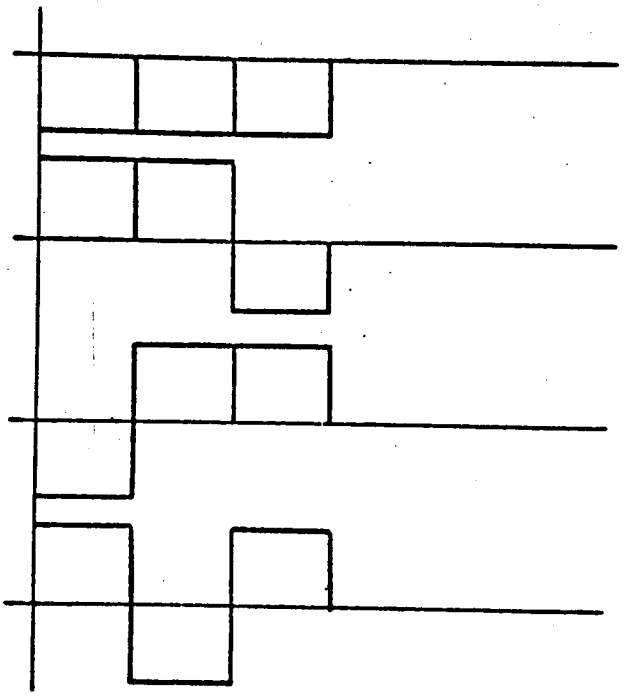


Figure 6.6 A Set of Waveforms achieving the Bound in section 4.8.

Before returning to the ranging problem and the application of PR-sequences to it, let us observe another quite interesting feature of these sequences: their ease of generation. The diagram

or figure 6.7 represents a set of three storage cells which retain a number, zero or one, until a clock pulse causes a shift. When the shift, or clock, pulse occurs, each binary digit shifts to the right, and the modulo-two sum of the contents of the last two cells is shifted into the first. By modulo-two sum we mean that

$$0 \oplus 0 = 0$$
$$0 \oplus 1 = 1$$
$$1 \oplus 0 = 1$$
$$1 \oplus 1 = 0$$

The reader may verify that the mechanism of Figure 6.7 called a shift register does indeed generate the complete PN-sequence given earlier of length seven regardless of which three binary digits are originally placed in the cells, unless all three digits are zero. (To be consistent with the earlier notation, the output "0" of this device must be converted to "-1".) In general a sequence of length $N = 2^n-1$ can be generated by an n-cell shift register.

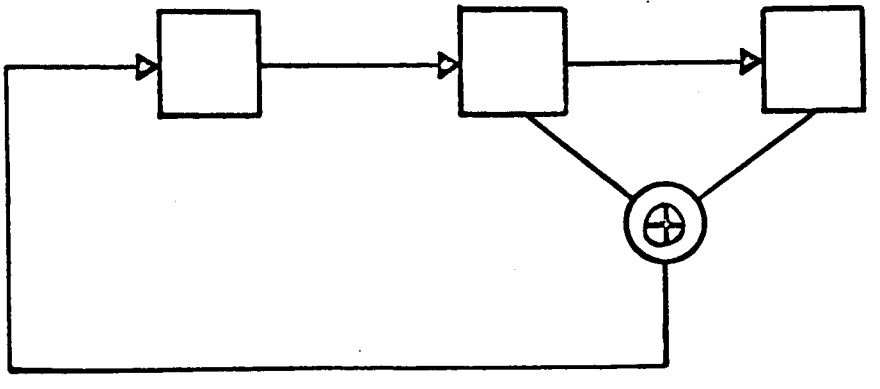


Figure 6.7 A three-stage shift register

6.3 Ranging with Pseudo-Random Sequences

The application of pseudo-random sequences to ranging should be fairly obvious from the discussion of the previous section. Since the

periodic autocorrelation of PN-sequences assumes the values

$$\phi(0) = N\Delta t$$

$$\phi(i\Delta t) = -\Delta t \qquad i = 1, 2, \ldots, N-1$$

The complete autocorrelation function is as shown in figure 6.8.

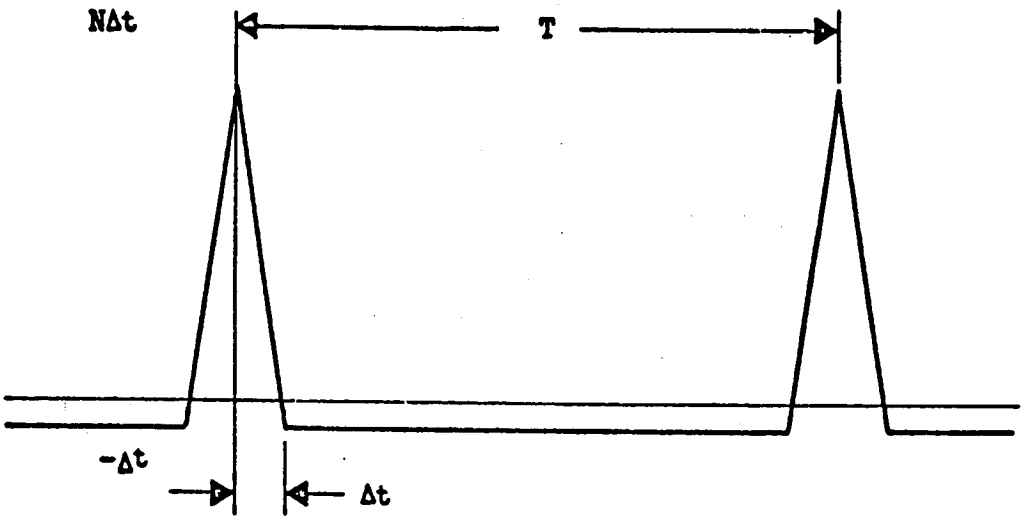


Figure 6.8 The PR sequence Autocorrelation Function

Suppose the m(t) in Equation (6.6) were this sequence rather than a single pulse of duration Δt. Then, since the optimum detector forms the integral

$$I'(\hat{t}_0) = \int_{\hat{t}_0}^{\hat{t}_0+T} m(t)\, y'(t)\, dt = \int_{\hat{t}_0}^{\hat{t}_0+T} x(t+\hat{t}_0) y'(t)\, dt \qquad (6.15)$$

where $y'(t) = A'x(t+t_0) + n(t)$

and x(t) represents the PR sequence, it follows that the output, in the absence of noise is

$$\left\{I'(\hat{t}_0)\right\}_{n(t)=0} = A' \int_{\hat{t}_0}^{\hat{t}_0+T} x(t+\hat{t}_0)\, x(t+t_0)\, dt = A'\phi(t_0-\hat{t}_0) \qquad (6.16)$$

$$= \begin{cases} N A' \left\{ \Delta t - \left(\frac{N+1}{N}\right) \left| t_0-\hat{t}_0 \right| \right\} & t_0-\Delta t < \hat{t}_0 < t_0 + \Delta t \\ -A'\,\Delta t & \text{otherwise} \end{cases}$$

Note the similarity between this equation and Equation (6.9). If the average power radiated in the two cases is equated, then $NA' = A$, and if N is large, as it must be for high range resolution, then $A' = A/N$ is negligible compared to A and the two cases are essentially equivalent. The PR sequence technique has the definite advantage that the same amount of power is always being radiated; the peak power is equal to the average power.

In ranging space probes some distance from the earth, it is difficult to reflect from them the amount of power necessary for high resolution ranging. However, the spacecraft can be equipped with a transponder which receives the transmitted signal, tracks both the carrier and the PR sequence, and regenerates them for retransmission back to earth. This increases the amount of power received at the earth tracking station by several orders of magnitude, and is, in fact, the technique most commonly used in conjunction with planetary missions.

It is well to observe at this point the close relationship between ranging and synchronization. The latter necessitates the establishment of a common time reference between the transmitter and the receiver, while ranging involves the measurement of the signal

transit time between the two. Both, therefore, rely upon the determination of the frequency and phase of a received signal. The frequency determination enables the receiver clock to be controlled so as to run at the same rate as the transmitter clock; the phase measurement, obtained, for example, by identifying the phase of a received PR sequence, establishes a common point in time. The PR sequence used in the ranging case, must be long enough to resolve the distance ambiguity. That is, if the sequence length corresponds to only 10 miles distance it may be difficult to resolve the ambiguity as to whether the vehicle being ranged is, for example, 200,010 miles away or 200,020 miles away. If, on the other hand its length corresponds to 100,000 miles, there would probably be little difficulty in determining whether the vehicle is 200,010 or 300,010 miles away. Similarly, in the synchronization problem, the sequence length must be great enough to resolve all time ambiguities not easily resolved otherwise.

6.4 Doppler Measurements

If the transmitter transmitting signals with a carrier frequency of f_c cycles per second, and the receiver are moving away from each other with a relative velocity of v meters/sec then in time T the distance between them has increased by an amount vT while f_cT carrier cycles have been transmitted. Designating the carrier wavelength by $\lambda = c/f_c$ where c is the velocity of light, we see that while f_cT cycles are transmitted, the distance has increased by vT/λ cycles and, hence, only $f_cT - vT/\lambda$ cycles are received. Thus, the apparent carrier frequency at the receiver is

$$f_c(1 - v/c)$$

The doppler frequency Δf, the difference between the transmitted frequency and the received frequency, is therefore

$$\Delta f = v/c \; f_c \quad . \tag{6.18}$$

If the transmitter and the receiver are moving towards each other with a velocity v, the doppler shift is, of course, the same in absolute magnitude but represents an apparent increase rather than a decrease in the received frequency.

In the case of the ranging of a space vehicle, in which the transmitted signal is either reflected back to the earth or received and re-transmitted to the earth by the vehicle, the doppler shift has been effected twice and hence yields a shift of 2v/c cps at the earthbound receiver. By measuring the received frequency, then, it is possible to measure the velocity of the spacecraft towards or away from the earth. This can be accomplished in a number of ways. Perhaps the most direct is to take the product of the transmitted signal carrier $\sqrt{2}$ sin $\omega_c t$ and the received signal carrier

$$\sqrt{2} \; A \sin\left[(\omega_c \pm 2\pi\Delta f)\, t+\phi\right]$$

obtaining

$$A \cos(2\pi\Delta f\, t \mp \phi) - A \cos\left[(2\omega_c \pm 2\pi\Delta f)\, t+\phi\right] \tag{6.19}$$

and to pass this signal through a bank of N very narrow bandpass filters, the i^{th} filter of which passes, unattenuated, all frequencies in the range

$$f_0 + \frac{i\delta f}{2} < f_0 + \frac{(i+1)\delta f}{2} \qquad i = 0, 1, \ldots N-1$$

and passes no energy outside this range.

(The frequency f_o ideally represents the lowest and $f_0 + N\delta f/2$ the highest expected doppler shift.) The filter which exhibits the greatest average output power then indicates the value of Δf and, thereby, the component of the velocity of the vehicle toward or away from the earth. If the sign of the frequency shift is in doubt, it can also be determined by some additional manipulation.

Since the carrier frequency is quite high, often on the order of 1000 megacycles/second, determining the doppler frequency to the nearest cycle/second, for example, establishes the velocity of the vehicle to within less than 1 mile/hour. Doppler measurements, when combined with some initial information concerning the vehicle position and velocity at some point in time, and with the knowledge of the trajectory which the vehicle must follow, provide a complete record of the position of the vehicle. The limitation to this approach in tracking a space vehicle lies in its dependence upon the need for rather precise measurements of the initial position and velocity data, upon the knowledge of the gravitational fields to which the vehicle is subject, and upon the mathematical difficulties in using all of this information to calculate the trajectories. Direct range measurements as described in the previous sections are also subject to the latter two limitations when it is desired to determine the trajectory of the vehicle. The doppler measurements and the direct range measurements do provide complementary information however, and the combination of these two measurements can often be used to improve the estimates of the physical parameters involved in the calculation of these trajectories.

6.5 Angle Measurement

As mentioned earlier, part of the knowledge of the position of the vehicle may depend upon the ability to measure the angular position of the antenna pointed directly at it; at any rate, it is necessary to be able to to keep the antenna accurately directed at the vehicle in order to realize its high gain. The coordinate system in which the antennas move depends, of course, upon the type of mount used to support them. The two most common coordinate systems, the azimuth-elevation (az-el) and the hour angle-declination (HA-DEC) systems are illustrated in figure 6.9. The solid line represents the direction in which the antenna is pointing. The planes of motion of the antenna can be visualized by keeping one of the coordinates fixed and moving the other of the same pair.

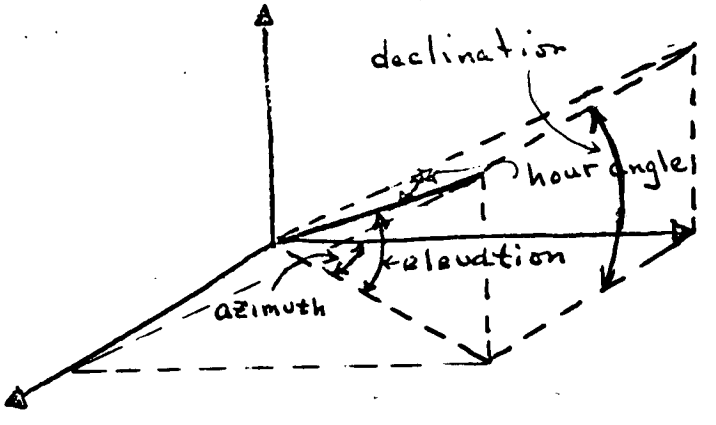


Figure 6.9 Two antenna coordinate systems

Since parabolic antennas are highly directional the coordinates can be varied until the received signal power is a maximum. While it is possible to point the antenna fairly precisely at the vehicle in this manner it is difficult to achieve the accuracy desired. Moreover, since the vehicle is generally moving in the antenna coordinate system,

it is desirable to provide a method whereby the antenna can automatically be kept aimed at the source . It is, therefore, useful to have some indication, from the received signal itself, as to the error in the antenna orientation, an indication which would enable this error to be decreased as well as to aid in tracking the moving vehicle.

This can be accomplished quite simply by placing at the antenna focus point, not one, but four pick-up points oriented as shown in figure 6.10. The signals from these four points are combined in three ways: (1) the difference between the sum of the outputs of the two "north" and that of the two "south" pickups and (2) that between the sum of the outputs of the two "east" and that of the two "west" pickups are formed as well as (3) the sum of the outputs all four pickups. If the antenna is properly oriented, the signals (1) and (2) will be zero.

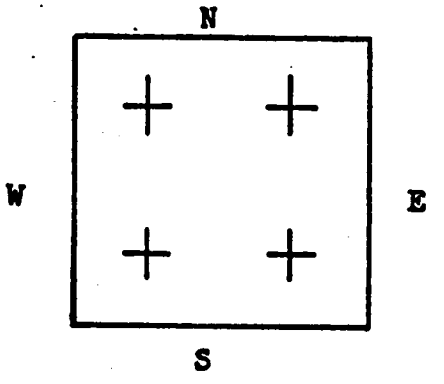


Figure 6.10. Antenna pick-up

If the antenna is pointed too far north, more signal power will be reflected towards the north pickups than to the south, thereby generating an error signal (1). This signal can be used to indicate the directional error causing the antenna to be shifted south until there is no more error signal. The same comments, of course, apply to orientation errors in the other directions.

Such a device is particularly useful during the initial stages of the flight when the spacecraft trajectory is not known precisely. In the later stages, the trajectory can usually be predicted with sufficient accuracy for the antenna to be controlled entirely by computer.

7 PIONEER, MARINER AND THE DSIF

In this concluding chapter we shall discuss briefly the communication systems which have actually been used in some of the lunar and deep-space programs, as well as the ground-based equipment known as the DSN.

7.1 The Pioneer Communications System

The Pioneer IV communications system consisted of three FM channels which were frequency multiplexed, the combined signal being used to phase modulate a carrier at 960 megacycles/sec.

Although there were only three channels, a total of seven different data sources were monitored through the use of time sharing and by a process of superimposing several signals simultaneously on one subcarrier and using the information redundancies to separate the data at the receiver. The first channel consisted of a temperature measurement. Because it was known that the temperature would not undergo any rapid changes, several "event indicators" were superimposed upon this signal. The events in question caused a jump in the temperature amplitude. These jumps could not be due to the temperature; they were recognized as the desired event indications and were subtracted out leaving only the temperature measurements. The several events to be designated in this way were: (1) an indication of the initiation of a "despinning" this way were: (1) an indication of the initiation of a "despinning" operation in which weights were extended to slow down the spinning of the probe; (2) a step voltage change indicating that the optical devices had received light due to the closeness to the moon; and (3) a series of steps (positive and negative, indicating that the optical cells were periodically seeing light and then darkness in

accordance with the spin-rate of the vehicle as it passed by the moon.

The second channel transmitted the integrated output of a Geiger counter. After the probe had passed through the Van Allen belts, and there was no more information to be measured, the input to this channel was switched to measure the output of the power amplifier of the transmitter.

Channel three was a method for transmitting the number of counts from a second Geiger-Mueller tube. Three signals were superimposed as shown in figure 7.1. The highest frequency output from the counter switched every 2^9 counts, the next every 2^{13} counts and the last every 2^{17} counts. These three signals were added as shown, and passed through a low-pass filter. As the counting rate increased, the high frequency component would be filtered out, and, hence, the number of counts would be scaled by a factor of 2^{13} rather than 2^9. As the number of counts per second increased even further, the middle frequency would also be filtered and the scale factor would be 2^{17}. The most significant digits of the count were thereby always transmitted while the bandwidth necessary to do this was kept relatively constant. This is an example of a self-adaptive data processing system.

The data signals from the three channels were applied to the input of three VCO's each centered at a different frequency to keep the outputs from overlapping in frequency. The signals were actually filtered first to limit the rate at which they could change levels as shown in Figure 7.2. This was done to prevent the VCO output signal from changing frequency too rapidly thereby enabling it to be coherently demodulated with a phase-locked loop. If the frequency input to the demodulating loop were to

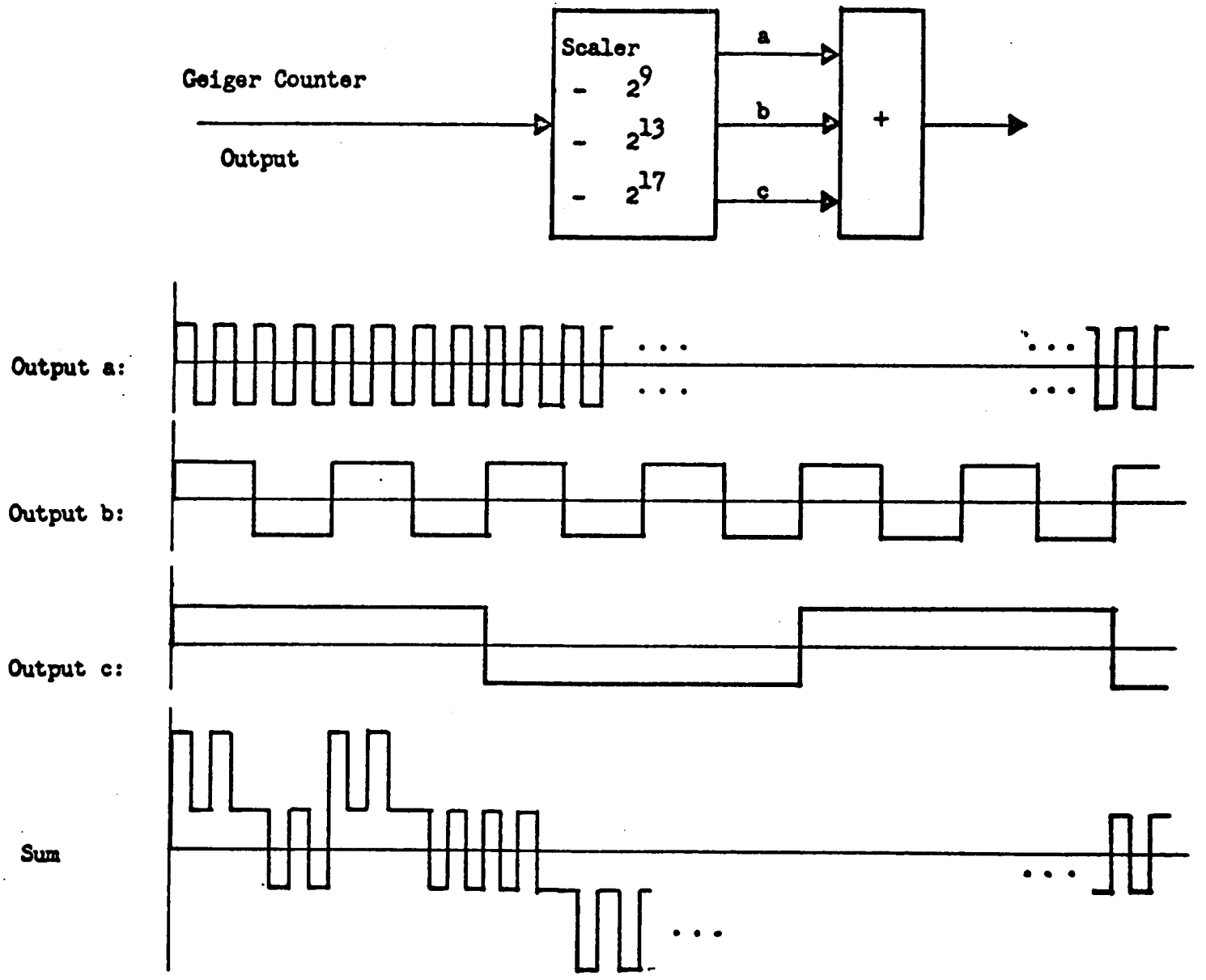
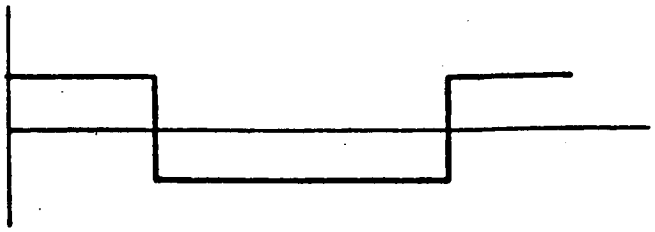
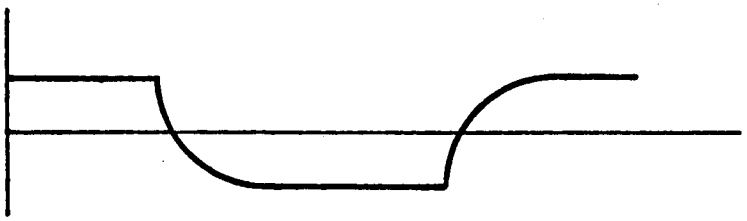


Figure 7.1 Channel three data signal



(a) Prefiltered data signal



(b) Filtered data signal

Figure 7.2. Data Signal Conditioning

change too abruptly the loop would go out of lock and the information would be lost until it could be again locked up. The three VCO ouptuts were added and used to phase modulate the carrier.

The total radiated power was 164 milliwatts, of which approximately 100 mw were in the carrier, 14 mw each in channels one and two, and 36 mw in channel three. The probe antenna was essentially a dipole antenna with a gain of 2.5 db.

The receiver antenna had an 85 foot diameter with a 40 db gain. The receiver effective noise temperature was 1630°K. Demodulation was accomplished with phase-locked loops as shown in Figure 7.3. The loop noise bandwidths were: carrier, 20 cps; channel one, 4 cps; channel two, 4 cps; channel three, 8 cps.

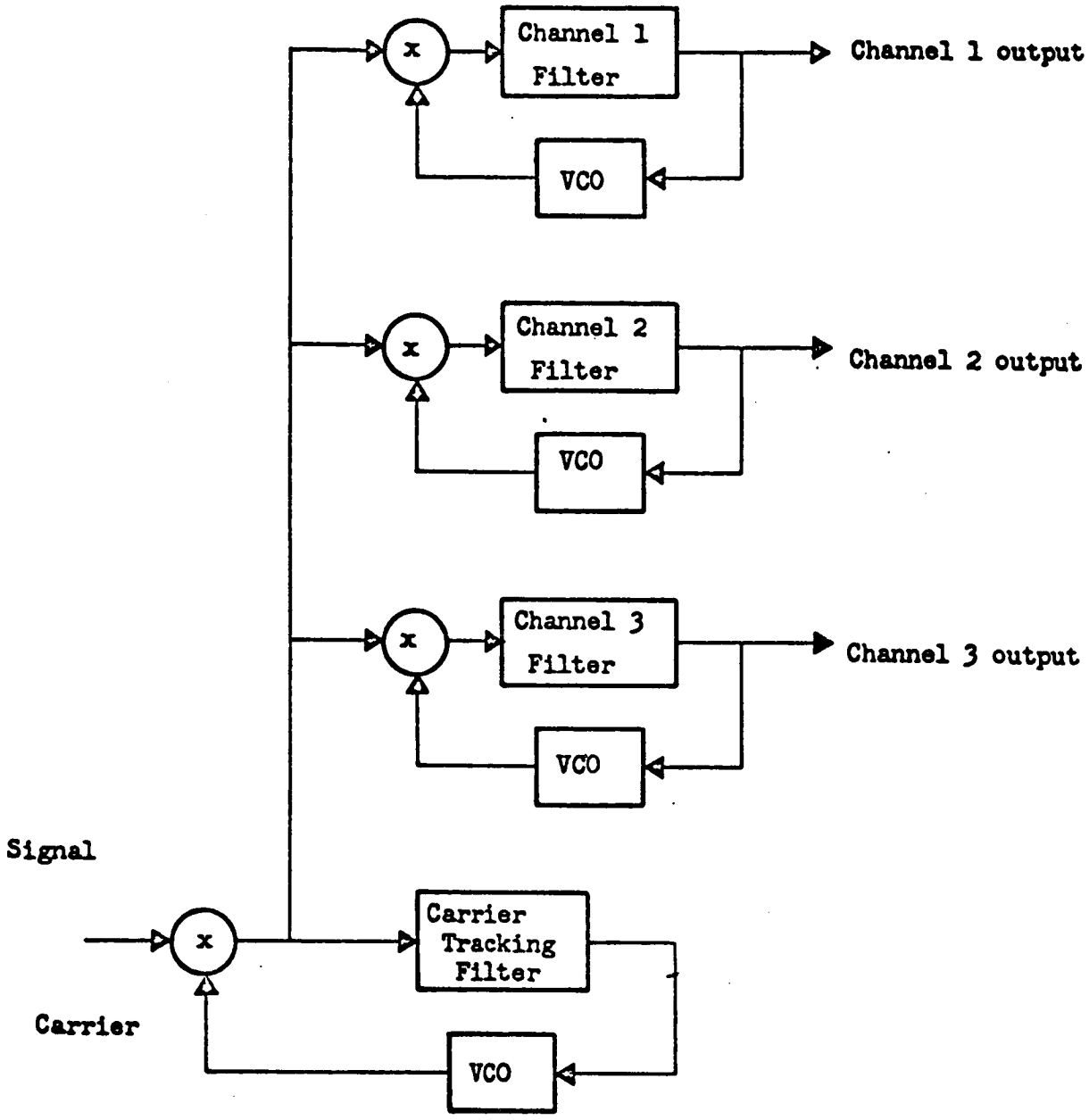


Figure 7.3 Pioneer Receiver Demodulator

Note that the carrier demodulation is accomplished by simply tracking the unmodulated carrier and taking as the partially demodulated signal the product of the tracked carrier with the received signal. This achieves the desired result only because the carrier is phase modulated with a small index of modulation. That is, let $x_1(t) + x_2(t) + x_3(t) = \phi(t)$, where $x_1(t)$ is the frequency modulated

signal from the i^{th} channel. Then the output of the phase modulator is

$$y(t) = \sqrt{2}\, A \sin\left[\omega_c t + \Delta\theta.\phi(t)\right]$$

$$= \sqrt{2}\, A \cos\left[\Delta\theta.\phi(t)\right] \sin \omega_c t + \sqrt{2}\, A \sin\left[\Delta\theta.\phi(t)\right] \cos \omega_c t \tag{7.1}$$

If $\Delta\theta$ is small enough then

$$\cos\left[\Delta\theta.\phi(t)\right] \approx 1$$

$$\sin\left[\Delta\theta.\phi(t)\right] \approx \Delta\theta.\phi(t)$$

and

$$y(t) \approx \sqrt{2}\, A \sin \omega_c t + \sqrt{2}\, A\, \Delta\theta.\phi(t) \cos \omega_c t \tag{7.2}$$

The carrier loop tracks the unmodulated component at the frequency $f_c = \frac{\omega_c}{2\pi}$. The output of the carrier loop VCO is therefore approximately $\sqrt{2} \cos \omega_c t$ and the product of this with the signal yields, neglecting the double frequency components,

$$A\Delta\theta.\phi(t) = A\Delta\theta \left[x_1(t) + x_2(t) + x_3(t)\right] \tag{7.3}$$

which is the desired frequency-multiplexed frequency modulated signal. Each of the three loops is designed to track only one of these three signals, the other two are eliminated by the tracking filter. That the FM signal is then demodulated follows from the discussion in section 3.6.

The Pioneer IV telecommunications system was then essentially a frequency multiplexed FM system which was coherently demodulated using phase-locked loops. No direct effort was made to remove the redundancy from the data signals, though, clearly, all three channels

produced very redundant information. Some effort was expended, however, to make use of the signal redundancy by superimposing several signals. This was, in some sense, a method of data compaction. In addition, the third channel did incorporate a simple but effective method of adapting to the data.

7.2 The Ranger Communication System

The major point of departure for the Ranger (nos. 6, 7, 8 and 9) communication system was in its capacity for televising back to earth photographs of the surface of the moon. The spacecraft contained six cameras. Two of these cameras, the full-scan (F) cameras, photographed a square segment of the moon each side covering 8.4°, while the four partial-scan (P) cameras covered an area of only 2.1° square. The video-tube and optical properties of the six cameras were essentially the same, but the partial scan cameras scanned only the central portion of the image. It was possible, therefore, to transmit more pictures (each, of course, covering a smaller range,) with the P cameras than with the F cameras. The relative areas covered by the six cameras is illustrated in figure 7.4 . The shutter speed of the P cameras was 2 milliseconds, while that of the F cameras was 5 m secs. The amount of time to transmit an image with the P and F cameras was 0.2 secs/image and 2.56 secs/image respectively. Each image was represented by 93-1/3 lines per degree of surface area photographed. The four partial-scan camera signals were time-multiplexed to form one telemetry channel and the two full-scan camera signals time-multiplexed to form a second. (For the last two minutes before impact both channels were used to transmit partial-scan pictures.) The bandwidth of each channel was 200 kilocycles/sec. The signal from each of these channels was pre-emphasized and used to FM modulate a carrier, one centered at

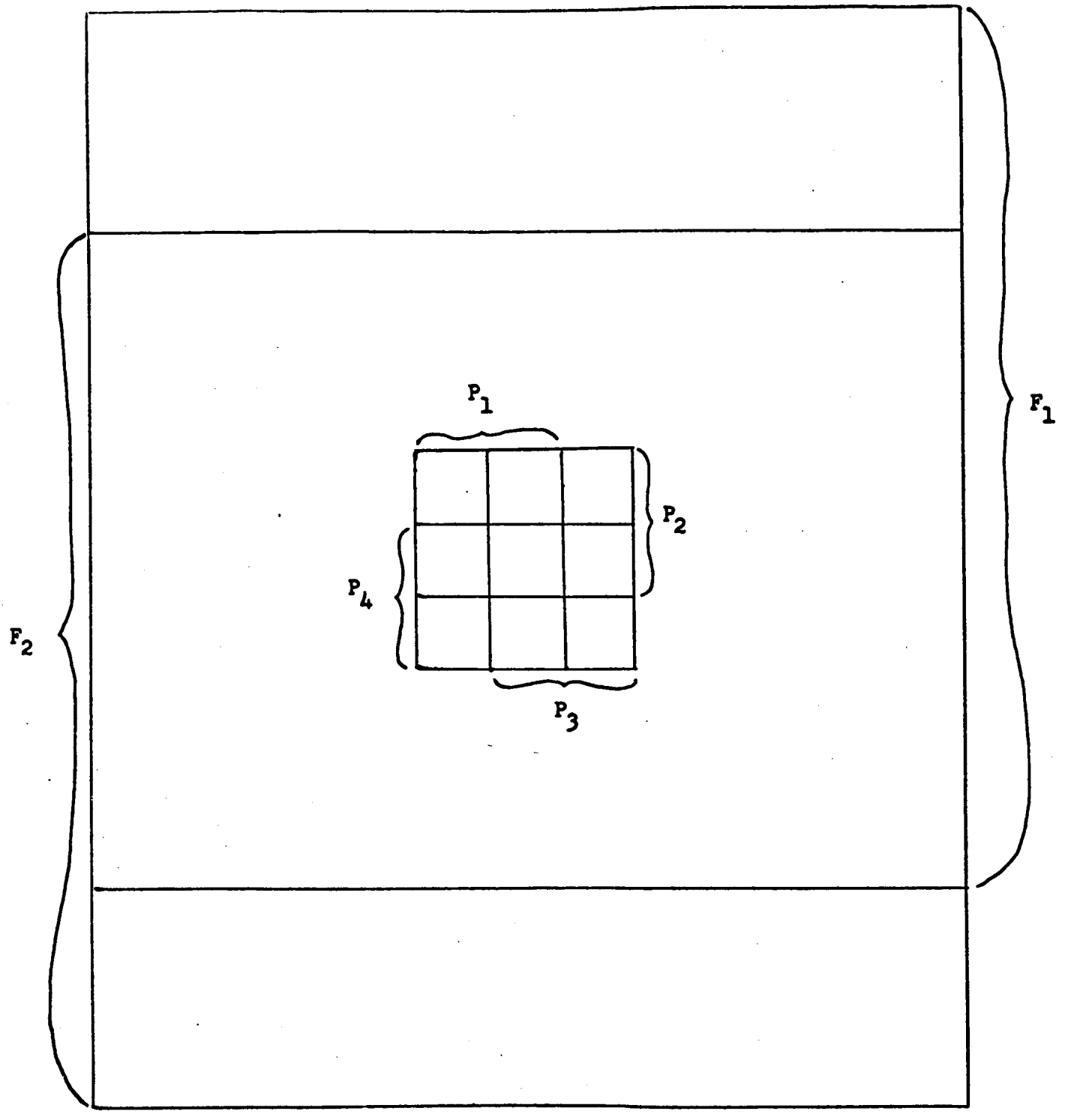


Fig. 7..4 The overlapping picture areas of the four partial-scan (P_1,P_2,P_3 and P_4) and two full-scan (F_1 & F_2) Ranger television cameras.

959.52 mc/sec and one at 960.58 mc/sec. The modulating signal was adjusted to provide a peak frequency deviation of 200 kc/sec. Thus, the modulation was relatively narrow-band FM. These two channels were frequency multiplexed with a third, a telemetry channel centered at 960.05 mc/sec and having an 80 kc/sec bandwidth, and transmitted over a 4' parabolic antenna. The transmitted power was 30 watts per television channel. The third telemetry channel conveyed some ninety engineering measurements taken during this terminal mode. This channel represented .3 watts of transmitted power. The data were sampled and used to PAM modulated a 225 cps carrier, which was in turn used to FM modulate a 960.05 mc/sec carrier.

The received signal was recorded on magnetic tape and after the two television channels were de-multiplexed (by filtering), demodulated (by an FM discriminator) and passed through a de-emphasis network, they were displayed on a cathode ray tube. The display was photographed both on 35mm and on Polaroid film. The third channel was demodulated by a phase-locked loop detector.

The preceeding paragraphs have briefly described the terminal mode of the communication system. The entire mission consisted of an injection mode, a cruise mode, a midcourse maneouver mode, and a terminal mode. During the first three of these modes the communication system was essentially that of the third (non-television) channel of the terminal mode. The transmitted power was three watts in the cruise and midcourse maneouver modes, and one watt in the injection mode. The difference between the modes was in the data sources monitored and in the antenna used (i.e. it was necessary to use an omni-directional antenna in the midcourse maneouver mode.) The different modes were

initiated by a signal from earth, although the sequencing in the terminal mode was automatic after it was once initiated.

7.3 The Mariner II Communication System (Venus probe, 1964)

The Mariner II telemetry system provides an interesting contrast to the Ranger system just discussed. The fundamental reason for the difference in the time systems was in the distances involved. Mariner was designed to communicate over distance of up to 50 million miles as opposed to the 1/4 to 1/2 million mile range required of the lunar probe. The data sources on board Mariner II were divided into engineering and scientific categories. The engineering sources included: measurements of the battery voltages, the solar-panel voltage, the earth-sensor temperature, roll, pitch, and yaw gyro indications, sun sensor outputs, propellant tank pressure and temperature, thermal shield temperature, etc. Some of the scientific experiments which had to be monitored included a microwave radiometer experiment; an infrared radiometer experiment, a magnometer experiment; charged particle flux experiments; solar plasma experiments, and a micrometeorite experiment.

To complicate the situation were two factors; (1) not all measurements were needed at all times; and (2) the telemetry capacity obviously decreased as the distance from the earth increased. It clearly would not be efficient to operate as if the worst condition were in effect; i.e., as if the spacecraft were at its maximum design distance from the earth and all measurements were to be transmitted. Fortunately many of the engineering measurements are most significant during the early powered stages of the flight and during the midcourse man while the telemetry capacity is relatively large and the scientific measurements are not of interest. The scientific measurements become most important during the fairly brief period of planetary encounter

when the spacecraft has approached its maximum design distance from the earth. For these reasons, the telemetry system was designed to operate in three successive modes: the launch mode; the cruise mode (after earth acquisition); and the planetary encounter mode (during which the scientific data was gathered and transmitted). Switching from one mode to another was accomplished by a command from the ground. Each telemetry mode had associated with it a certain subset of the total measurements which were to be monitored.

The various data sources were time multiplexed, each of the signals being observed continually for a fixed length of time T_W. A clock, generating $R_W = \frac{1}{T_W}$ pulses per second switched the different measurements into the output. Since some measurements needed to be observed more frequently than others, a process of commutation evolved so that it was possible to sample some sources every tenth time, some every one hundredth time, and some only every one thousandth time. The technique for accomplishing this is indicated in figure 7.4. The switches marked S indicate some of the various possible mode switches. The scientific data was transmitted during the encounter mode. The ten-input one-output boxes indicate commutators which read the input from the i^{th} switch until it receives a pulse and then switches to the $i + 1^{st}$ switch. After it reaches the tenth position it starts over again at the first. At the output the information was sampled once every T_W seconds and converted to binary PCM form. Each sample was represented by words of seven bits each. The bit clock therefore generated $R_B = 7\ R_W$ pulses per second; the bit interval was $T_B = 1/R_B$ seconds.

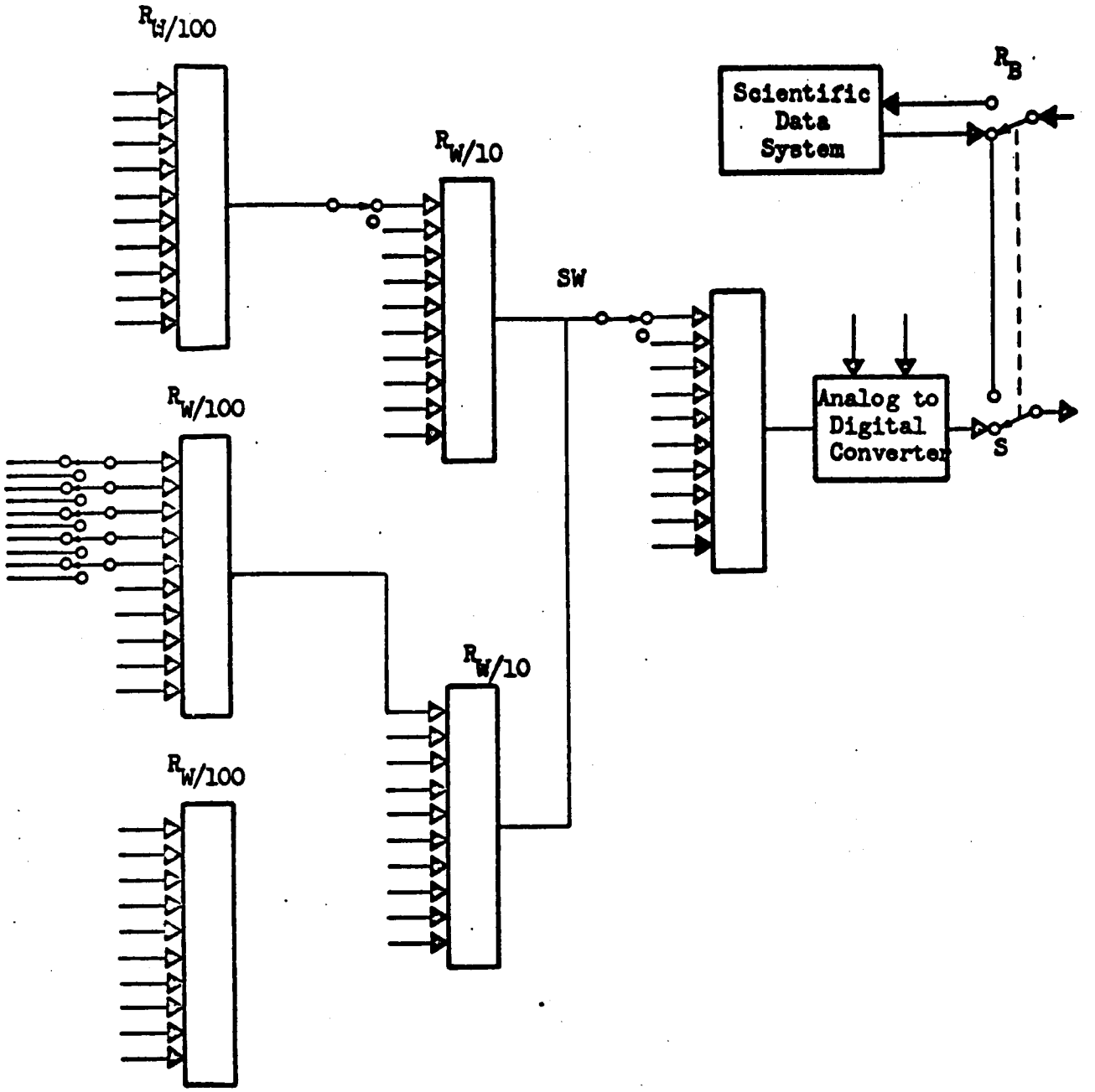


Figure 7.5 Telemetry Commutator and Mode Switching Mechanization

The input to the modulator was a sequence of PCM binary bits. At the output of the modulator a one was represented by the signal $x_1(t) = \sqrt{2}$ A sin $\omega_s t$ and a zero by $x_2(t) = -\sqrt{2}$ A sin $\omega_s t$. Since the two signals represented equal power, and since

$$\rho_{12} = \frac{\int_0^T x_1(t)\, x_2(t) dt}{\left[\int_0^T x_1^2(t)\, dt \int_0^T x_2^2(t)\, dt\right]^{\frac{1}{2}}} = -1 \qquad (7.4)$$

this corresponds to a two-level biorthogonal code. The associated bit-error probabilities are as shown in Figure 4.11, for $n_2 = \log_2 N_2 = 1$.

A convenient alternative way of representing the modulator output is by the waveform

$$\sqrt{2}\ A \cos\left(\omega_s t + \pi/2\, m(t)\right) \qquad (7.5)$$

where $m(t) = +1$ or -1 for $nT_B < t < (n+1)T_B$ depending upon whether the bit is a one or a zero. This is therefore an example of discrete phase-shift keying (PSK).

This PSK signal is in turn used to phase-modulate the 960 mc/sec carrier. Essentially the same carrier demodulator described in Section 7.1 for the Pioneer system was used here. Again because of the non-linear aspects of coherent PM demodulation, it is required that about 50% of the power be left in the carrier; i.e., that the modulation index be kept relatively small.

A block diagram of the essential components of the telemetry demodulator is shown in Figure 7.6. Note that the optimum demodulation scheme

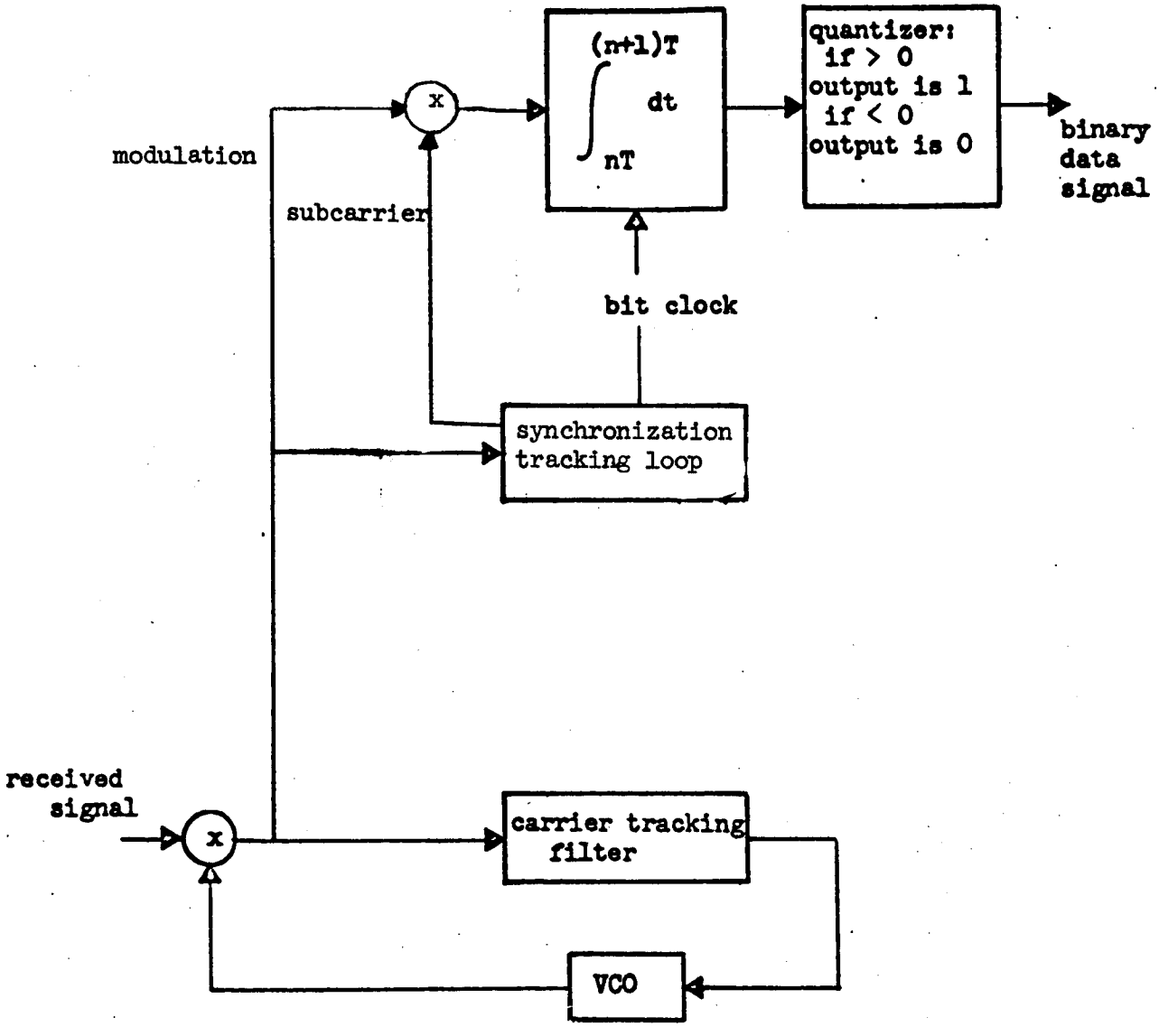


Figure 7.6 The Mariner Telemetry Demodulator

as discussed in Chapter Four is used to convert the signal $-A\, m(t) \sin\theta$ to a sequence of binary bits. This, of course, demands the knowledge of the instants of time when one bit ends and the next one begins. This bit synchronization as well as word and frame synchronization must be provided as discussed in Section 4.12. The technique by which this information was obtained in the Mariner II System is entirely analogous to the PR ranging system discussed in Section 6.3. As pointed out there, ranging and synchronization are strongly related problems; in this case essentially the same solution was used for synchronization as is often used for ranging. A PR sequence containing 63 bits with period $T_B/9$ was transmitted. Since $63\ T_B/9 = 7\ T_B = T_W$, the period of the PR sequence was equal to the word interval. By identifying a particular phase of the PR sequence with the beginning of a word, word synchronization, and a fortiori, bit synchronization, is obtained as soon as the phase of the PR sequence is detected. Moreover, since the PR-sequence symbol period and the subcarrier periods were integrally related, this same information was also used to generate the local replica of the subcarrier needed for the product demodulator. The PR bit period was chosen to be $T_B/9$ rather than T_B in order to increase the resolution, i.e., in order to decrease the region of uncertainty Δt (cf. Section 6.3).

Frame synchronization, the knowledge of that instant of time in which the complete cycle of data observations begins again, was provided by making the first word of each frame consist of ones only and prohibiting the all-ones word from occurring elsewhere. (If the data signal were truly all ones it would be replaced by the signal 1 1 1 1 1 1 0.) Consequently, the occurrence of the all-ones word indicated the beginning of a new frame.

In summary, then, the Mariner telecommunications system might

be classified as a PCM/PSK/PM system. The data is time multiplexed, sampled and converted to binary PCM. The binary bits are used to PSK modulate a subcarrier which in turn phase modulates a carrier. The purpose of a subcarrier is to move the data spectrum away from the zero frequency range so that, when this signal is used to modulate the carrier, the receiver carrier loop is able to track the carrier without interference from the modulation. Approximately 50% of the total power must be left in the carrier. This is not necessarily a limitation, however, since unmodulated carrier power is essential both for locking up the receiver loop and for tracking the vehicle. The Mariner transmitting antenna was parabolic with a diameter of four feet. The radiated power was 3 watts; the telemetry rate was $33\frac{1}{3}$ bits/sec. in the launch mode, and $8\frac{1}{3}$ bits/sec. in the other two modes of operation.

The command link (ground-to-spacecraft) was quite similar to the telecommunications link. Several important differences should be mentioned however. First, while it is important to have reliable communication between the spacecraft and ground, it is imperative to have extremely high reliability in the other direction. Commands from the earth are used to alter the mode of operation and even the trajectory of the spacecraft. An error in the reception of a command could easily defeat the purpose of the entire mission. Although much more power (up to 10kw) is radiated from the ground than from the spacecraft, the ground-to-vehicle link is not significantly better than the vehicle-to-ground link. This is because the vehicle receiver effective noise temperature (about 7300°K in the case of Mariner II) is considerably higher than that of the ground receiver, and, in addition, the vehicle receiver antenna gain

is less than its transmitter antenna gain. This latter factor is because commands must be received regardless of the attitude of the vehicle; the receiver antenna should be omni-directional and therefore has ideally a gain of 0 db. Should something go wrong with the vehicle attitude control, for example, the mission from that point on might be a total failure were it not possible to communicate with the probe, regardless of its orientation. The reliability of the ground-to-vehicle communication is increased, however, by transmitting the data at a very slow bit rate. As seen in Figure 4.11, the error probability is a rapidly decreasing function of the parameter ST_B/N_0 where T_B is the time spent per bit. Since relatively few commands need to be transmitted T_B can be made quite long and the bit-error probability satisfactorily small.

7.4 The Mariner IV Telemetry System (Mars probe, 1965)

The Mariner IV space probe was designed to approach within a distance of between 4,000 and 20,000 miles from the planet Mars. Scientific measurements were to be taken en route and in the vicinity of Mars, and television pictures of the surface of Mars to be taken during the nearest approach to the planet. The communication system was basically that of Mariner II with some significant modifications. Although Mariner IV was designed to transmit information over nearly three times as far as did the Mariner II probe, the basic bit rate was kept the same ($8\frac{1}{3}$ bps). The increased distance was compensated for by going to higher frequencies (2116 megacycles/sec as opposed to 960 mc/sec) and by increasing the spacecraft transmitter power from 3 to 10 watts.

The communication system was designed to operate in four different modes:

(1) Launch, initial acquisition, and cruise mode. Blocks of engineering data 140 bits long are alternated with scientific data blocks of length 280 bits.

(2) Maneuver mode: Only engineering data is transmitted. This mode was to be used at other times to aid in failure analysis, if necessary.

(3) Planetary Encounter Mode: This mode was to be used during the $13\frac{1}{3}$ hours of the nearest approach to Mars. Only science data is sampled and transmitted.

(4) Post-encounter mode. Video and engineering measurements that were recorded during the encounter phase are transmitted.

The transmission rates were $33\frac{1}{3}$ bps through the first midcourse maneuver and $8\frac{1}{3}$ bps thereafter. The engineering measurements were essentially those of Mariner II. The scientific measurements included, in addition to the television pictures;

(1) A cosmic-ray telescope to measure the proton and alpha-particle densities within a specified cone angle.

(2) A cosmic dust detector to determine the momenta and mass distribution of dust particles uncountered, and their direction of travel, relative to the vehicle.

(3) A trapped radiation detector to measure cosmic rays and energetic electrons in space, and, in particular, to detect the presence of magnetically trapped particles in the vicinity of Mars.

(4) An ion chamber to measure the flux of all energetic particles capable of penetrating its walls.

(5) A plasma probe to measure the flux of the positively charged components of the solar plasma.

(6) A magnetometer to measure the magnitude and direction of the interplanetary magnetic field and to determine whether a magnetic field exists around Mars.

(7) Ultraviolet photometers to measure hydrogen (1216 Å) and oxygen (1300 Å) radiation originating in the outer atmosphere of Mars.

This latter system, and, of course, the television system operated only during encounter. They were sighted toward Mars by a planetary scanning system which tracked the illuminated planet.

The television camera took approximately 20 photographs of the surface of Mars over a period of about 20 minutes. The picture consisted of 200 lines each with 200 picture elements quantized to 64 levels. Each picture was represented by 240,000 binary bits which were recorded on an endless-loop tape recorder at the rate of 10,700 bits/sec. The tape recorder capacity was 5.28 $\times 10^6$ bits enabling it to record up to 22 complete pictures. This information was transmitted to Earth during the post-encounter mode. Each picture transmission lasted approximately 8 hours. Between each picture 1.5 hours of engineering data was transmitted.

The spacecraft contained two antennas, a low-gain (6db) antenna giving essentially uniform coverage in the earth-directed hemisphere of the probe, and a high-gain (23db) parabolic antenna, with an eliptical cross-section (21" minor axis, 46" major axis.) The spacecraft telemetry system was automatically designed to switch from the high-to the low-gain antenna if signals were not periodically received from Earth to prohibit this action. This was to insure that if some malfunction caused the high-gain antenna not to be pointed towards

Earth, at least it would be possible to communicate with the probe through the low-gain antenna. At encounter, were it necessary to use the low-gain antenna, only the Goldstone tracking station would have the necessary power (100 kw) to transmit commands to the spacecraft. The Mariner IV receiver noise temperature was approximately 2700°K.

Before concluding this chapter we will discuss briefly the ground based part of the communication system called the Deep Space Net.

7.5 The Deep Space Net

The Deep Space Net consists of three basic elements: The Deep Space Instrumentation Facility (DSIF), the Space Flight Operations Facility (SFOF), and the Ground Communications System (GCS). The DSIF is comprised of ten transmitting-receiving stations suitably located so as to enable 24-hour communication with any deep-space probe. Four of these stations (the Venus, Pioneer, Echo, and Mars sites) are located at Goldstone, California; one at Cape Kennedy, Florida; one on Ascension Island, one at Madrid, Spain, one at Johannesburg, South Africa, and one each at Canberra (Tidbinbilla) and Woomera, Australia. The facilities at all these sites are similar. The Venus site at Goldstone, however, is used as a research and development facility for testing new equipment before it is installed at the rest of the sites. A 210 ft steerable parabolic antenna is in operation at the Mars site while the other sites, with the exception of the Cape Kennedy and Ascension sites, have 85 ft antennas. The latter two stations are used only for the early minutes of the launch and, consequently, do not require large antennas, since the received power is much greater at this stage of the flight. Moreover, since they must be capable of following a relatively rapidly moving object, their smaller mass is a definite asset. The major sites are equipped with helium-cooled ruby maser amplifiers. The overall receiver noise

temperature is between 30° and 60°K including background noise.

The SFOF is located at the Jet Propulsion Laboratory in Pasadena. It is the control center of the DSN system, to which the information is transmitted from each of the DSIF sites. This information is used to evaluate the progress of the flight, to determine the flight parameters, and to make decisions regarding commands to be sent to the spacecraft, etc. To handle and process this information the SFOF is equipped with one of the largest digital computing facilities ever assembled.

The GCS is the designation for the system to convey information from the various DSIF sites to the SFOF. Through it each station is connected to the DSIF through various teletype lines.

BIBLIOGRAPHY

CHAPTER IV

Section One

Davenport, Wilbur B., Jr., and William L. Root, An Introduction to the Theory of Random Signals and Noise, McGraw-Hill Book Co., Inc., New York, 1958.

Rice, Steven O., "Mathematical Analysis of Random Noise," Bell System Technical Journal, 23, pp. 282-332, 1944; 24 pp. 46-156, 1945.

Section Two

Blackwell, Lawrence A., and Kenneth L. Kotzebue, Semiconductor-Diode Parametric Amplifiers, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1961.

Hogg, Christopher A., and Lawrence G. Sucsy, Masers and Lasers, Maser and Laser Associates, Cambridge 40, Mass., 1962.

Fradin, C. Z., Microwave Antennas, Pergamon Press, London, 1961.

Hansen, R. D., ed., Microwave Scanning Antennas, Vol. I., Academic Press, New York, 1964.

Pierce, John R., Electrons, Waves and Messages, Hanover House, Garden City, New York, 1956.

Section Three

Black, Harold S., Modulation Theory, D. Van Nostrand Co., Inc., New York, 1953.

Schwartz, Mischa, Information Transmission, Modulation and Noise, McGraw-Hill Book Co., Inc., New York, 1959.

Jaffe, R., and E. Rechtin, "Design and Performance of Phase-Lock Circuits Capable of Near-Optimum Performance over a Wide Range of Input Signal and Noise Levels," IRE Trans. on Information Theory, March 1955, pp. 66-76.

Section Four

Fano, Robert M., Transmission of Information, MIT Press & John Wiley and Sons, Inc., New York, 1961.

Golomb, S. W., et al., Digital Communications, Prentice-Hall, Englewood Cliffs, New Jersey, 1964

Section Five

Elias, Peter, "Predictive Coding," IRE Trans, on Information Theory, March 1955, pp. 16-33.

Huffman, D. A., "A Method for the Construction of Minimum Redundancy Codes," Proc. IRE, Vol. 40, p. 1098, Sept., 1952.

Section Six

Golomb, S. W., et al., Digital Communications, Prentice-Hall, Englewood Cliffs, New Jersey, 1964.

Section Seven

Martin, Benn. D., "The Pioneer IV Lunar Probe: A Minimum Power FM/PM System Design," JPL Technical Report No. 32-225, March 15, 1962.

Riddle, F. M., et al., "JPL Contributions to the 1962 National Telemetering Conference," JPL Technical Memorandum No. 33-88, May 21, 1962.

Mathison, Richard P., "Mariner Mars 1964 Telemetry and Command System," Sixth Winter Convention on Military Electronics, Feb., 1965, Los Angeles, California.